

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2249	65.0	2404	1	US-08-311-023-3	Sequence 3, Appli
2	1288	37.2	1310	1	US-08-311-023-1	Sequence 1, Appli
3	360	10.4	385	4	US-09-712-016-78	Sequence 78, Appli
c 4	240	6.9	241	4	US-09-389-6261-360	Sequence 360, App
c 5	240	6.9	241	4	US-09-620-405B-360	Sequence 360, App
c 6	240	6.9	241	4	US-09-433-826B-360	Sequence 360, App
c 7	240	6.9	241	4	US-09-604-287A-360	Sequence 360, App
c 8	165	4.8	177	3	US-09-020-956-162	Sequence 162, App
c 9	165	4.8	177	3	US-09-030-607-162	Sequence 162, App
c 10	165	4.8	177	4	US-09-439-313-162	Sequence 162, App
c 11	165	4.8	177	4	US-09-332-616A-162	Sequence 162, App
c 12	165	4.8	177	4	US-09-233-149A-162	Sequence 162, App
c 13	110.6	3.2	1888	3	US-09-599-360B-38	Sequence 38, Appli
c 14	59.8	1.7	285	3	US-08-990-571-64	Sequence 64, Appli
c 15	59.8	1.7	285	4	US-09-528-784A-64	Sequence 64, Appli
c 16	59.8	1.7	285	4	US-09-569-098A-64	Sequence 64, Appli
c 17	59.8	1.7	342	3	US-08-990-571-65	Sequence 65, Appli
c 18	59.8	1.7	342	4	US-09-528-784A-65	Sequence 65, Appli
c 19	59.8	1.7	342	4	US-09-569-098A-65	Sequence 65, Appli
c 20	59.8	1.7	351	3	US-08-990-571-60	Sequence 60, Appli
c 21	59.8	1.7	351	4	US-09-528-784A-60	Sequence 60, Appli
c 22	59.8	1.7	351	4	US-09-569-098A-60	Sequence 60, Appli
c 23	59.8	1.7	356	3	US-08-990-571-63	Sequence 63, Appli
c 24	59.8	1.7	356	4	US-09-528-784A-63	Sequence 63, Appli
c 25	59.8	1.7	356	4	US-09-569-098A-63	Sequence 63, Appli
c 26	59.8	1.7	358	3	US-08-990-571-58	Sequence 58, Appli
c 27	59.8	1.7	358	4	US-09-528-784A-58	Sequence 58, Appli

Db 1141 GAAATGCTGCACAAATGATGCTAGTGACCATGGATGAGCGCTGGGGGTATTTCTTTT 1200
 QY 2319 CAGATGCTGGGATGCTTTGGGTTTGGGAATATATGTTACTTATTTCCATATTTGAACAT 2378
 Db 1201 CAGATGCTGGGATGCTTTGGGTTTGGGAATATATGTTACTTATTTCCATATTTGAACAT 1259
 QY 2379 AAATGCTGTTTGGTATATAATTTCTAGTTAAAGGTTTAAATGCTAGAGTAGCT 2430
 Db 1260 AAATGCTG-TTGGTATATAATTTCTAGTTAAAGGTTTAAATGCTAGAGTAGCT 1310

RESULT 3

US-09-712-016-78
 ; Sequence 78, Application US/09712016
 ; Patent No. 6468790
 ; GENERAL INFORMATION:
 ; APPLICANT: Giese, Klaus
 ; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER
 ; FILE REFERENCE: 200130.460
 ; CURRENT APPLICATION NUMBER: US/09/712,016
 ; CURRENT FILING DATE: 2000-11-13
 ; PRIOR APPLICATION NUMBER: US/09/417,615
 ; PRIOR FILING DATE: PRIOR FILING DATE: 1999-10-13
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 78
 ; LENGTH: 385
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 US-09-712-016-78

Query Match 10.4%; Score 360; DB 4; Length 385;
 Best Local Similarity 98.2%; Pred. No. 1.3e-90;
 Matches 375; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2453 TAGGTCATGAGGAGATGAGTTGTTGATGCTGCTACTATCCACGCTTAAAGTAGTGGGTTT 2512
 Db 1 TCGGTCATGAGGAGA--GATTTGATGCTGCTACTATCCACGCTTAAAGTAGTGGGTTT 58

QY 2513 TGTGATTTTCTATGATTAATGCTGCTGTTTACAAAGTCAGTTAAAGTAGCTTTTAAAT 2572
 Db 59 TGTGATTTTCTATGATTAATGCTGCTGTTTACAAAGTCAGTTAAAGTAGCTTTTAAAT 118

QY 2573 ATTTAAGTTATTCCTGCGATAAAATCTGTATGCAATTCACCGTATTACCACT 2632
 Db 119 ATTTAAGTTATTCCTGCGATAAAATCTGTATGCAATTCACCGTATTACCACT 178

QY 2633 TTATATGTAACAAGAGATTTGGCATGACATGTTCTGTATGTTTCAGGGAAAAATGCT 2692
 Db 179 TTATATGTAACAAGAGATTTGGCATGACATGTTCTGTATGTTTCAGGGAAAAATGCT 238

QY 2693 TTAATGCTTTTCAAGAACTAACACAGTTATTCCTATCTGATTTAGGCTCTGAAGA 2752
 Db 239 TTAATGCTTTTCAAGAACTAACACAGTTATTCCTATCTGATTTAGGCTCTGAAGA 298

QY 2753 ACTGCTGGTGTAGGAATAAGATGTGCATGAGCCTAAAATACCAAGAAAGCTTATAC 2812
 Db 299 ACTGCTGGTGTAGGAATAAGATGTGCATGAGCCTAAAATACCAAGAAAGCTTATAC 358

QY 2813 TGAATTTAAGCAAGAAATAAA 2834
 Db 359 TGAATTTAAGCAAGAAATAAA 380

RESULT 4

US-09-389-681-360/c
 ; Sequence 360, Application US/09389681A
 ; Patent No. 6518237
 ; GENERAL INFORMATION:
 ; APPLICANT: Yuqiu, Jiang
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
 ; FILE REFERENCE: 210121.470C3
 ; CURRENT APPLICATION NUMBER: US/09/389,681A
 ; CURRENT FILING DATE: 1999-09-02
 ; NUMBER OF SEQ ID NOS: 463
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 360
 ; LENGTH: 241
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(241)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-389-681-360

Query Match 6.9%; Score 240; DB 4; Length 241;
 Best Local Similarity 100.0%; Pred. No. 3.1e-57;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2720 TTATTCCTATACCTGATTTTAGTCTCTGGAAGTCTGCTGTTTAGAATAAGATGT 2779
 Db 241 TTATTCCTATACCTGATTTTAGTCTCTGGAAGTCTGCTGTTTAGAATAAGATGT 182

QY 2780 GCATGAAGCTCTAAATACCAAGAAAGCTTATCTAGATTTAAAGCAAGAAATAAGGAGA 2839
 Db 181 GCATGAAGCTCTAAATACCAAGAAAGCTTATCTAGATTTAAAGCAAGAAATAAGGAGA 122

QY 2840 AAAGAGAAAGATCTGAGAAATTTGGGAGGAGCATAGATTTCTTATAAAATCAGAAATTTGTT 2899
 Db 121 AAAGAGAAAGATCTGAGAAATTTGGGAGGAGCATAGATTTCTTATAAAATCAGAAATTTGTT 62

QY 2900 GTAAATAGAGGGAGAGAAATTTAGAAATTAAGTATAAAAGGAGAAATTAAGTATAGTAC 2959
 Db 61 GTAAATAGAGGGAGAGAAATTTAGAAATTAAGTATAAAAGGAGAAATTAAGTATAGTAC 2

RESULT 5

US-09-620-405B-360/c
 ; Sequence 360, Application US/09620405B
 ; Patent No. 6528054
 ; GENERAL INFORMATION:
 ; APPLICANT: Jiang, Yuguu
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Hepler, William T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE REFERENCE: 210121.470C3
 ; CURRENT APPLICATION NUMBER: US/09/620,405B
 ; CURRENT FILING DATE: 2000-07-20
 ; NUMBER OF SEQ ID NOS: 495
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 360
 ; LENGTH: 241
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(241)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-620-405B-360

Query Match 6.9%; Score 240; DB 4; Length 241;
 Best Local Similarity 100.0%; Pred. No. 3.1e-57;
 Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2720 TTATTCCTATACCTGATTTTAGTCTCTGGAAGTCTGCTGTTTAGAATAAGATGT 2779
 Db 241 TTATTCCTATACCTGATTTTAGTCTCTGGAAGTCTGCTGTTTAGAATAAGATGT 182

Db 241 TTATCTCTACTGGATTAGTCTCTGAGAACTGCTGGTGTCTTAGGAATAGATGT 182
QY 2780 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 2839
Db 181 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 122
QY 2840 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 2899
Db 121 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 62
QY 2900 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2959
Db 61 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2

RESULT 6

US-09-433-826B-360/c
; Sequence 360, Application US/09433826B
; Patent No. 6579973

GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqui

; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND

; FILE REFERENCE: 210121.470C4

; CURRENT APPLICATION NUMBER: US/09/433,826B

; CURRENT FILING DATE: 1999-11-03

; NUMBER OF SEQ ID NOS: 474

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 360

; LENGTH: 241

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(241)

; OTHER INFORMATION: n = A,T,C or G

US-09-433-826B-360

Query Match 6.9%; Score 240; DB 4; Length 241;

Best Local Similarity 100.0%; Pred. No. 3.1e-57;

Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2720 TTATCTCTACTGGATTAGTCTCTGAGAACTGCTGGTGTCTTAGGAATAGATGT 2779

Db 241 TTATCTCTACTGGATTAGTCTCTGAGAACTGCTGGTGTCTTAGGAATAGATGT 182

QY 2780 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 2839

Db 181 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 122

QY 2840 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 2899

Db 121 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 62

QY 2900 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2959

Db 61 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2

RESULT 7

US-09-604-287A-360/c

; Sequence 360, Application US/09604287A

; Patent No. 6366572

GENERAL INFORMATION:

; APPLICANT: Jiang, Yuqui

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Xu, Jiangchun

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Hepler, William T.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.470C7

; CURRENT APPLICATION NUMBER: US/09/604,287A

; CURRENT FILING DATE: 2000-06-22

; NUMBER OF SEQ ID NOS: 489

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 360

; LENGTH: 241

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(241)

; OTHER INFORMATION: n = A,T,C or G

US-09-604-287A-360

Query Match 6.9%; Score 240; DB 4; Length 241;

Best Local Similarity 100.0%; Pred. No. 3.1e-57;

Matches 240; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2720 TTATCTCTACTGGATTAGTCTCTGAGAACTGCTGGTGTCTTAGGAATAGATGT 2779

Db 241 TTATCTCTACTGGATTAGTCTCTGAGAACTGCTGGTGTCTTAGGAATAGATGT 182

QY 2780 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 2839

Db 181 GCATGAAGCTTAAATACCAAGAAGCTTATAGTGAATTAAGCAAAAGAAATAAAGGAGA 122

QY 2840 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 2899

Db 121 AAGAGAGAAGTCTGAGAAATGGGAGGCAATAGATCTTATAAAATACAAAATTTGTT 62

QY 2900 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2959

Db 61 GTAAATAGAGGGGAGAAATTTAGAAATTAAGTATAAAAGGAGCAATTAAGATAGATAC 2

RESULT 8

US-09-020-956-162/c

; Sequence 162, Application US/09020956

; Patent No. 6261562

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHOD:

; NUMBER OF SEQUENCES: 178

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/020,956

; FILING DATE: 09-FEB-1998

CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Makl, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.427C2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (206) 622-4900

; TELEFAX: (206) 682-6031

; INFORMATION FOR SEQ ID NO: 162:

; SEQUENCE CHARACTERISTICS:

```

; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
;
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
;
; FILE REFERENCE: 210121.427C8
;
; CURRENT APPLICATION NUMBER: US/09/352,616A

```

```
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-352-616A-162

Query Match
Best Local Similarity 4.8%; Score 165; DB 4; Length 177;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3159 TTATGATATACCAAGCTGGTTATGCGCAAGTATATATACCAAAAAGCTGTAGACTG 3218
DB 165 TTATGATATACCAAGCTGGTTATGCGCAAGTATATATACCAAAAAGCTGTAGACTG 106

QY 3219 GATGTTCTGGTTACTGTTTACAAAATATACAGAGTAGTAAACTTTGATATATAGAG 3278
DB 105 GATGTTCTGGTTACTGTTTACAAAATATACAGAGTAGTAAACTTTGATATATAGAG 46

QY 3279 GATATATAAATACACTAAGTATATGATGATTCAGTACAGAAAGT 3323
DB 45 GATATATAAATACACTAAGTATATGATGATTCAGTACAGAAAGT 1

RESULT 12
US-09-232-149A-162/c
; Sequence 162, Application US/09232149A
; Patent No. 6465611
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; FILE OF INVENTION: CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C6
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 162
; LENGTH: 177
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-162

Query Match
Best Local Similarity 4.8%; Score 165; DB 4; Length 177;
Matches 165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3159 TTATGATATACCAAGCTGGTTATGCGCAAGTATATATACCAAAAAGCTGTAGACTG 3218
DB 165 TTATGATATACCAAGCTGGTTATGCGCAAGTATATATACCAAAAAGCTGTAGACTG 106

QY 3219 GATGTTCTGGTTACTGTTTACAAAATATACAGAGTAGTAAACTTTGATATATAGAG 3278
DB 105 GATGTTCTGGTTACTGTTTACAAAATATACAGAGTAGTAAACTTTGATATATAGAG 46

QY 3279 GATATATAAATACACTAAGTATATGATGATTCAGTACAGAAAGT 3323
DB 45 GATATATAAATACACTAAGTATATGATGATTCAGTACAGAAAGT 1

RESULT 13
US-09-599-360B-38
; Sequence 38, Application US/09599360B
; Patent No. 6548633
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Bouquelieret, L.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: Complementary DNA's Encoding Proteins with Signal Peptides
```

```
; FILE REFERENCE: GENSET.050CP3
; CURRENT APPLICATION NUMBER: US/09/599,360B
; CURRENT FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: 60/113,686
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/141,032
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 09/469,099
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent.pm
; SEQ ID NO 38
; LENGTH: 1888
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 139..1389
; NAME/KEY: sig_peptide
; LOCATION: 139..198
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 5.00
; OTHER INFORMATION: seq HLLAGFCVWVVG/VV
; NAME/KEY: polva_signal
; LOCATION: 1854..1859
; NAME/KEY: polyA_site
; LOCATION: 1873..1888
US-09-599-360B-38

Query Match
Best Local Similarity 3.2%; Score 110.6; DB 4; Length 1888;
Matches 211; Conservative 0; Mismatches 149; Indels 1; Gaps 1;

QY 1908 GCGCTCGCCACTTTGGCCCTGGATGAATGGGTGATGGCTGCACAAATTCAGCGAT 1967
DB 1273 GGTGGCACTGATATACAGTGGTGGTCTCTCTGGAGATGGTCTACACAACTCACTGAT 1332

QY 1968 GCGCTAGCAATTTGGTGGCTTTTACTCAAGGCTATCAAGTGGTTAAGTACTTCGT 2027
DB 1333 GGGCTGGCCATPAGGTGCTGCTCTCTCTGATGGCTTC-CGGCGGCTCAGTACACCTTA 1391

QY 2028 GCTGTGTTCTGTCATGAGTTGCTCATGAATAGTGGTGGTGGTGGTGGTGGTGGTGGT 2087
DB 1392 GCGGTCTCTCCATGAGCTGCCACCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1451

QY 2088 GCGATGACCGTTAAGCAGGCTGCTCTTATATATGATGATGATGATGATGATGATGAT 2147
DB 1452 GGGCTGCTCTTTCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511

QY 2148 GGAATGGCAACAGGAATTTTCATGCTATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2207
DB 1512 GGTGCGAGTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1571

QY 2208 GCACTTACTGCTGGCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 2267
DB 1572 GGGGCTACTGCTGGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1631

QY 2268 C 2268
DB 1632 C 1632

RESULT 14
US-08-990-571-64/c
; Sequence 64, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
```

CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,571
FILING DATE: 11-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 201021.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-990-571-64

Query Match	1.7%;	Score 59.8;	DB 3;	Length 285;
Best Local Similarity	65.2%;	Pred. No. 7e-07;		
Matches	88;	Conservative 0;	Mismatches 47;	Indels 0; Gaps 0;
QY	405	ATCATATACACCATTGCCAGCACCATCACTCAGCCAGCGACTCATCTCAGGCATGAG	464	
Dd	192	AACCAAGTTCGACTAGGGCCCCACGCTTTCATAGGCCACCACGCTTCACTAGGCCCCACG	133	
QY	465	CGTCACTCAGACCATGACATCACTCAGACCCAGCAGCATCATCTTGACCATGATCATCAC	524	
Dd	132	CTTCCTAGGCCACCACGCTTCATAGGCCCCCACGCTTCACTAGGCCCCACGACTTCAC	73	
QY	525	TCTCACATAAATCAT	539	
Dd	72	TAGGCCAACCGACTT	58	

```

RESULT 15
US-09-528-784A-64/C
; Sequence 64, Application US/09528784A
; Patent No. 6451315
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C4
; CURRENT APPLICATION NUMBER: US/09/528,784A
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Babesia microti
US-09-528-784A-64

```

Query Match 1.7%; Score 59.8; DB 4; Length 285;
Best Local Similarity 65.2%; Pred. No. 7e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QV 405 ATCCATTACACATGACCAGCACCATCACTCAGACCGAGCATCACTCAGACCATGAG 464

		192	AACAGGTTCCACTAGGCCCCACAGCTTCACTAGGCCCCACAGCTTCACTAGGCCCCACAG	133
Dδ				
Qγ		465	GCTCCTCAGACCATGAGCATCACTCAGACAGCAGCAGCATCACTCTGACCATGATCATCAC	524
Dδ		132	CTTCACTAGGCCCCACAGCTTCACTAGGCCCCACAGCTTCACTAGGCCCCACAGCTTCACT	73
Qγ		525	TCTCACCATTATCAT	539
Dδ		72	TAGGCCCCACAGCTT	58

Search completed: September 12, 2003, 16:55:40
Job time : 197.714 secs

Result No.	Score	Query	Length	DB	ID	Description
1	3335	96.4	3586	14	US-10-176-847-51	Sequence 51, Appl
2	3335.6	96.4	3537	14	US-10-198-846-10007	Sequence 10007, A
3	2716	78.5	2744	14	US-10-177-581-170	Sequence 170, App
4	2716	78.5	2744	14	US-10-177-293-263	Sequence 263, App
5	1141.4	33.0	1193	10	US-09-925-300-591	Sequence 591, App
6	900.4	36.0	1449	14	US-10-198-846-10321	Sequence 10321, A
7	694.2	20.1	1032	14	US-10-198-846-13805	Sequence 13805, A
8	578.4	16.7	606	14	US-10-198-846-8490	Sequence 8490, Ap
9	578.4	16.7	1598	14	US-10-198-846-11215	Sequence 11215, A
10	404.8	11.7	483	10	US-09-867-701-2021	Sequence 2021, Ap
c 11	390.8	11.3	395	9	US-09-778-320-94	Sequence 94, Appl
c 12	390.8	11.3	395	9	US-09-910-689-94	Sequence 94, Appl
c 13	390.8	11.3	395	13	US-10-910-742-94	Sequence 94, Appl
14	389.6	11.3	504	11	US-09-918-993-22272	Sequence 22272, A
15	360	10.4	385	9	US-09-827-669-78	Sequence 78, Appl
16	360	10.4	385	12	US-10-236-006-78	Sequence 78, Appl

[illegible]

Db 2505 TTCTAGTAAAGTTAAAGCTAGAGTAGCTTAAAAAGTTGTCATAGTTTCAGTAGGTCA 2564
Qy |||||
Db 2460 TAGGAGATGAGTTGTATGCTACTATGACAGCGTTTAAAGTTAGTGGTTTGTGATT 2519
Qy |||||
Db 2565 TAGGAGATGAGTTGTATGCTACTATGACAGCGTTTAAAGTTAGTGGTTTGTGATT 2624
Qy |||||
Db 2520 TTGTATGAATATGCTGCTGTACAAAGTCAGTTAAAGGTACGTTTAAATATTAAG 2579
Qy |||||
Db 2625 TTGTATGAATATGCTGCTGTACAAAGTCAGTTAAAGGTACGTTTAAATATTAAG 2684
Qy |||||
Db 2580 TTATCTATCTTTGAGATAAATCTGTATGTGCAATTCACGGTATTAACAGTTTATAT 2639
Qy |||||
Db 2685 TTATCTATCTTTGAGATAAATCTGTATGTGCAATTCACGGTATTAACAGTTTATAT 2744
Qy |||||
Db 2640 GTAACAAGATTTGGCATGACGATGCTGTGATGTTTACAGGAAATATGCTTTAAGC 2699
Qy |||||
Db 2745 GTAACAAGATTTGGCATGACGATGCTGTGATGTTTACAGGAAATATGCTTTAAGC 2804
Qy |||||
Db 2700 TTTTTCAGAACTAACACAGTTATCCCTATCTGATTTAGCTCTCTGAAGATGCTG 2759
Qy |||||
Db 2805 TTTTTCAGAACTAACACAGTTATCCCTATCTGATTTAGCTCTCTGAAGATGCTG 2864
Qy |||||
Db 2760 GTTCTTAGGAATAAGATGTGATGAGCTTAAATACCAAGAAAGCTTTATCTGAATTT 2819
Qy |||||
Db 2865 GTTCTTAGGAATAAGATGTGATGAGCTTAAATACCAAGAAAGCTTTATCTGAATTT 2924
Qy |||||
Db 2820 AAGCAAAAGATTAAGGAGAAAGAGAGAGATCTGAGAAATGGGGAGGAGATGCTTTA 2879
Qy |||||
Db 2925 AAGCAAAAGATTAAGGAGAAAGAGAGAGATCTGAGAAATGGGGAGGAGATGCTTTA 2984
Qy |||||
Db 2880 TAAATACACAAATTTTGTAAATAGAGGGAGAAATTTAGAAATTAAGTATAAAAAG 2939
Qy |||||
Db 2985 TAAATACACAAATTTTGTAAATAGAGGGAGAAATTTAGAAATTAAGTATAAAAAG 3044
Qy |||||
Db 2940 GAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2999
Qy |||||
Db 3045 GAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3104
Qy |||||
Db 3000 CTTAGTGAGCTCTCTAT 3059
Qy |||||
Db 3105 CTTAGTGAGCTCTCTAT 3164
Qy |||||
Db 3060 TAAATATATTTAATGATTAACGATATACACTTGACCAAGAAATTTGGAATTTCAAAAT 3119
Qy |||||
Db 3165 TAAATATATTTAATGATTAACGATATACACTTGACCAAGAAATTTGGAATTTCAAAAT 3224
Qy |||||
Db 3120 GTTCGTGGGGTTATATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3179
Qy |||||
Db 3225 GTTCGTGGGGTTATATACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 3284
Qy |||||
Db 3180 TTATTTGCCAAGTTAT 3239
Qy |||||
Db 3285 TTATTTGCCAAGTTAT 3344
Qy |||||
Db 3240 ACAAAATATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3299
Qy |||||
Db 3345 ACAAAATATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3404
Qy |||||
Db 3300 ATCATTTGATTCGATTCAGAAAGTACTTTGATATCTCTCAGTGTCTCAGTGTCTCAGT 3359
Qy |||||
Db 3405 ATCATTTGATTCGATTCAGAAAGTACTTTGATATCTCTCAGTGTCTCAGTGTCTCAGT 3464
Qy |||||
Db 3360 TGAGCAATTTCTTTAT 3419
Qy |||||
Db 3465 TGAGCAATTTCTTTAT 3524
Qy |||||
Db 3420 AGATGTTCTTTTATACAAATAAATTCCTTATATACAGCTTG 3461
Qy |||||
Db 3525 AGATGTTCTTTTATACAAATAAATTCCTTATATACAGCTTG 3566

US-10-198-846-10007
; Sequence 10007, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10007
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3533, 3534, 3535, 3536, 3537
; OTHER INFORMATION: n = A, T, C or G
US-10-198-846-10007
Query Match 96.44; Score 3335.6; DB 14; Length 3537;
Best Local Similarity 99.24; Pred. No. 0;
Matches 3433; Conservative 0; Mismatches 4; Indels 25; Gaps 7;
Qy 3 CGTCCGGAATTCGCGACGAGCGCGTTCGCGCTGTAGAGATTTCTCGACACACC 62
Db |||||
Qy 3 CGTCCGGAATTCGCGACGAGCGCGTTCGCGCTGTAGAGATTTCTCGACACACC 62
Db |||||
Qy 63 AGTGGCCCGTGTGGACAAACCTGCGCGGTGGCGGCGCGCGCGCGCGCGCGCG 122
Db |||||
Qy 63 AGTGGCCCGTGTGGACAAACCTGCGCGGTGGCGGCGCGCGCGCGCGCGCGCG 122
Db |||||
Qy 123 GGAGAGAGCGGCAATGCGGAGGAGTATCTGTATCTTGATCTGACCTTGGCCCTC 182
Db |||||
Qy 123 GGAGAGAGCGGCAATGCGGAGGAGTATCTGTATCTTGATCTGACCTTGGCCCTC 182
Db |||||
Qy 183 TCTGTCAAAATCCCTTCATGAACCTTAAAGCAGCTCTTCCCGACACACTGAGAAA 242
Db |||||
Qy 183 TCTGTCAAAATCCCTTCATGAACCTTAAAGCAGCTCTTCCCGACACACTGAGAAA 242
Db |||||
Qy 243 ATTAGTCGGAATTTGGGAATCTGGCAATTAATGTGACCTTGGCAATTTCCACAGGCAATAT 302
Db |||||
Qy 243 ATTAGTCGGAATTTGGGAATCTGGCAATTAATGTGACCTTGGCAATTTCCACAGGCAATAT 302
Db |||||
Qy 303 CATCTACACAGCTTTTACCGCTATGAGAAATTAATCTTTGTCAGTTCAAGGGTTC 362
Db |||||
Qy 303 CATCTACACAGCTTTTACCGCTATGAGAAATTAATCTTTGTCAGTTCAAGGGTTC 362
Db |||||
Qy 363 AGAAAATTTACTTCAAAATATAGCATAGATTAAGATTAAGAAATTCATATACCATGAC 422
Db |||||
Qy 363 AGAAAATTTACTTCAAAATATAGCATAGATTAAGATTAAGAAATTCATATACCATGAC 422
Db |||||
Qy 423 CAGGACATCTACTCAGACACGAGCATCAGTACAGACCATGAGGTCTACTCAGACCATGAG 482
Db |||||
Qy 423 CAGGACATCTACTCAGACACGAGCATCAGTACAGACCATGAGGTCTACTCAGACCATGAG 482
Db |||||
Qy 483 CATCACTCAGACACGAGCATCAGTCTGACCATCATCATCTCTCACCATAATCATGCT 542
Db |||||
Qy 483 CATCACTCAGACACGAGCATCAGTCTGACCATCATCATCTCTCACCATAATCATGCT 542
Db |||||
Qy 543 GCTTCTGTTAAATAAGCGAAAAGCTTTTGGCCAGACCATGACTCAGATAGTTTCAGGT 602
Db |||||
Qy 525 GCTTCTGTTAAATAAGCGAAAAGCTTTTGGCCAGACCATGACTCAGATAGTTTCAGGT 584
Db |||||
Qy 603 AAAGATCTTAGAAACAGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGTGAGA 662

1081 CTCGAAAGACCTATTCATTAACAAATAGCCTGGGTGGTGTATAGCCATTTCCATCA 1140
1063 CTCGAAAGACCTATTCATTAACAAATAGCCTGGGTGGTGTATAGCCATTTCCATCA 1122
1141 TCAGTTTCCTGCTCTGCTGGGGGTATCTTATAGTCCCTCATGAATGGGTGGTGTTC 1200
1123 TCAGTTTCCTGCTCTGCTGGGGGTATCTTATAGTCCCTCATGAATGGGTGGTGTTC 1182
1201 AATTTCTCCTGAGTTTCTTGTGGCACTGGCCGTTGGACTTTGAGTGGTGTATGCTTTT 1260
1183 AATTTCTCCTGAGTTTCTTGTGGCACTGGCCGTTGGACTTTGAGTGGTGTATGCTTTT 1242
1261 TACACCTTCTTCCACATCTCATGCAAGTACACCACTAGCTCATGCAATGAGAACAG 1320
1243 TACACCTTCTTCCACATCTCATGCAAGTACACCACTAGCTCATGCAATGAGAACAG 1302
1321 CAATGGAATGAAAGAGGACCACTTTTCAGTCACTCTCTCTCCTCAAAACATAGAAGAA 1380
1303 CAATGGAATGAAAGAGGACCACTTTTCAGTCACTCTCTCTCCTCAAAACATAGAAGAA 1362
1381 GTGCTATTTGATTCACGTGGAAGGTCTAAGAGTCTAGAGGCTGTATTTCAATGT 1440
1363 GTGCTATTTGATTCACGTGGAAGGTCTAAGAGTCTAGAGGCTGTATTTCAATGT 1422
1441 TTCTTTGTAACATGCTCTCACAATTCATCAACAAATTTAAAGATAGAAGAAAAGATC 1500
1423 TTCTTTGTAACATGCTCTCACAATTCATCAACAAATTTAAAGATAGAAGAAAAGATC 1482
1501 AGAAGAACCTGAAATGATGATGATGGAGATTAAGAGCAGTTGTCCAGATGATGAT 1560
1483 AGAAGAACCTGAAATGATGATGATGGAGATTAAGAGCAGTTGTCCAGATGATGAT 1542
1561 CTCACCTTCAACAAATGAGGAGAAAGATGATGATGAGATGATGATGATGATGATGAT 1620
1543 CTCACCTTCAACAAATGAGGAGAAAGATGATGATGAGATGATGATGATGATGATGAT 1602
1621 GAGCAGACTCAAGAGCCCTCCCACTTTGATTTCTCAGCAGCCTGCAGCTTTGGAGAG 1680
1603 GAGCAGACTCAAGAGCCCTCCCACTTTGATTTCTCAGCAGCCTGCAGCTTTGGAGAG 1662
1681 AGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1663 AGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722
1741 GGTGCAAGAAATGAAATGCAATTCACATTCACAGATACATCGGCCAGTCAGACGATCA 1800
1723 GGTGCAAGAAATGAAATGCAATTCACATTCACAGATACATCGGCCAGTCAGACGATCA 1782
1801 TTCACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1783 TTCACACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1842
1861 CTCACAGTCACAGCCAGCGTACTCTCGGAGGAGCTGAAAGATGCGGCGTCGCCACTT 1920
1843 CTCACAGTCACAGCCAGCGTACTCTCGGAGGAGCTGAAAGATGCGGCGTCGCCACTT 1902
1921 TGGCCTGGATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1903 TGGCCTGGATGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962
1981 GTGCTGCTTTTACTGAAGCTTATCAGTGGTGTAAAGTACTCTGCTGCTGCTGCTGCTGCTGCT 2040
1963 GTGCTGCTTTTACTGAAGCTTATCAGTGGTGTAAAGTACTCTGCTGCTGCTGCTGCTGCTGCT 2022
2041 ATGAGTTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2023 ATGAGTTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2082
2101 AGCAGGCTGCTCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2083 AGCAGGCTGCTCTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2142
2161 GAATTTTCATTTGGTCATTAAGCTGAAATGTTTCTATGATGGATATTTGCACTTACGCTG 2220

2143 GAATTTTCATTTGGTCATTAAGCTGAAATGTTTCTATGGAATTTTGCCTTACCTGCTG 2202
2221 GCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
2203 GCTTATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2262
2281 GTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2263 GTGCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2322
2341 GTTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2400
2323 GTTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2382
2401 TCTAGTTAAAGTTTAAATGCTAGAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGT 2460
2383 TCTAGTTAAAGTTTAAATGCTAGAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGT 2442
2461 AGGAGAT 2520
2443 AGGAGAT 2502
2521 TTGATTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
2503 TTGATTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2562
2581 TATTTCTATCTTGAGATATAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
2563 TATTTCTATCTTGAGATATAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2622
2641 TAAACAGAGATTTGGCATGACATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
2623 TAAACAGAGATTTGGCATGACATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2682
2701 TTTTCCAGACTTAACACATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760
2683 TTTTCCAGACTTAACACATGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2742
2761 TG 2762
2743 TG 2744

RESULT 4

US-10-177-293-263

; Sequence 263, Application US/10177293

; Publication No. US20030124128A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Glatt, Karen

; APPLICANT: Zhao, Xumei

; APPLICANT: Gannavarpu, Manjula

; APPLICANT: Kamatkar, Shubhangi

; APPLICANT: Mertens, Maureen

; APPLICANT: Myer, Vic

; APPLICANT: Wang, Youzhen

; APPLICANT: Xu, Yongyao

; APPLICANT: Hoersch, Sebastian

; APPLICANT: Monahan, John

; APPLICANT: Meyers, Rachel E.

; APPLICANT: East Jr., Robert C.

; APPLICANT: Hortobagyi, Gabriel N.

; APPLICANT: Pusztai, Lajos

; APPLICANT: Meric, Funda

; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.

; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER

; FILE OF INVENTION: MRI-038

; CURRENT APPLICATION NUMBER: US/10/177,293

; FILING DATE: 2002-06-21

; PRIOR APPLICATION NUMBER: US 60/299,887

; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-263

Query Match 78.5%; Score 2716; DB 14; Length 2744;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2744; Conservative 0; Mismatches 0; Indels 18; Gaps 1;
QY 1 CTGTCGCGAATTCGGCAGACGCGGCTGTCGCGCTGTGAGAGATTTCTCGAAGACA 60
DB 1 CTGTCGCGAATTCGGCAGACGCGGCTGTCGCGCTGTGAGAGATTTCTCGAAGACA 60
QY 61 CCAGTGGGCGCTGTGGACAAACACCTGGCGCGTGGCGCGTGGGACACAGAGGC 120
DB 61 CCAGTGGGCGCTGTGGACAAACACCTGGCGCGTGGCGCGTGGGACACAGAGGC 120
QY 121 GCGGAGACGAAGGCGCAATGCGAGGAATATCTGTAATCTTGATCTGACCTTGCCC 180
DB 121 GCGGAGACGAAGGCGCAATGCGAGGAATATCTGTAATCTTGATCTGACCTTGCCC 180
QY 181 TCTCTGTCACAAATCCCTTCATGAACATAAAGCAGCTGCTTCCCGCAGACCACTGAGA 240
DB 181 TCTCTGTCACAAATCCCTTCATGAACATAAAGCAGCTGCTTCCCGCAGACCACTGAGA 240
QY 241 AATATAGTCGAATGGGAATCTGGCAATTAATGTTGACCTGGCAATTTCCACAGGCAAT 300
DB 241 AATATAGTCGAATGGGAATCTGGCAATTAATGTTGACCTGGCAATTTCCACAGGCAAT 300
QY 301 ATCATCTACACAGCTTTCTACCGCTATGGAGAAATTAATTTCTGAGTTGAAGGT 360
DB 301 ATCATCTACACAGCTTTCTACCGCTATGGAGAAATTAATTTCTGAGTTGAAGGT 360
QY 361 TCAGAAATTAATCTCAAAATATAGGCATAGATTAAGATTAAGAAATCCATATACCATG 420
DB 361 TCAGAAATTAATCTCAAAATATAGGCATAGATTAAGATTAAGAAATCCATATACCATG 420
QY 421 ACCAGACCATCACTCAGACCAACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATG 480
DB 421 ACCAGACCATCACTCAGACCAACGAGCATCACTCAGACCATGAGCGTCACTCAGACCATG 480
QY 481 AGCATCACTCAGACCAACGAGCATCACTCAGACCATGATCATCACTCAGACCAATCAATG 540
DB 481 AGCATCACTCAGACCAACGAGCATCACTCAGACCATGATCATCACTCAGACCAATCAATG 540
QY 541 CTGCTTCTGGTAAATTAAGCGAAAGAGCTTTTGGCCAGACCATGACATGATGTTGAG 600
DB 541 CTGCTTCTGGTAAATTAAGCGAAAGAGCTTTTGGCCAGACCATGACATGATGTTGAG 600
QY 601 GTAAAGATCTAGAAACAGCGGGAAGAGCTCAGCGACCAAGCAATGCCAGTGSTA 660
DB 601 GTAAAGATCTAGAAACAGCGGGAAGAGCTCAGCGACCAAGCAATGCCAGTGSTA 660
QY 661 GAAGATGTCGAAGACAGTGTAGTCTAGTGAAGTGAAGTCACTGCTGATGATGATGATG 720
DB 661 GAAGATGTCGAAGACAGTGTAGTCTAGTGAAGTGAAGTCACTGCTGATGATGATGATG 720
QY 721 TCTCTGAGGACTCACTTTCTAGAGCAATAGAGACTCCAGAGCCTGGAAATCTTCC 780
DB 721 TCTCTGAGGACTCACTTTCTAGAGCAATAGAGACTCCAGAGCCTGGAAATCTTCC 780

DB 703 TCTCTGAGGAACTCACCTTCTAGAGACAATAGAGACTCCAGACTCGAAATCTCTCC 762
QY 781 CCAAAGATGAAGAGCTCCACTCCACCAGTGTACATCAAGAGCGGGTGAAGCGGC 840
DB 763 CCAAAGATGAAGAGCTCCACTCCACCAGTGTACATCAAGAGCGGGTGAAGCGGC 822
QY 841 TGGCTGTAGGAAACAAATGAATCTGTGAGTGAGCCCGGAAAGGCTTTATGATTTCA 900
DB 823 TGGCTGTAGGAAACAAATGAATCTGTGAGTGAGCCCGGAAAGGCTTTATGATTTCA 882
QY 901 GAAACACAAATGAAATCTCTCAGAGTGTTCATGATCAATCAAGCTGACATCTCATG 960
DB 883 GAAACACAAATGAAATCTCTCAGAGTGTTCATGATCAATCAAGCTGACATCTCATG 942
QY 961 GCATGGGCATCCAGGTTCCGCTGAATCAACAGAGTTCACATCTCTGTCGACCATCA 1020
DB 943 GCATGGGCATCCAGGTTCCGCTGAATCAACAGAGTTCACATCTCTGTCGACCATCA 1002
QY 1021 TCAACAAATGATCTAGATCTTCTGATTCATACAGTGAAGAGAGCTGAATCC 1080
DB 1003 TCAACAAATGATCTAGATCTTCTGATTCATACAGTGAAGAGAGCTGAATCC 1062
QY 1081 CTCGAAAGACCTTATTCATCAAAATAGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1140
DB 1063 CTCGAAAGACCTTATTCATCAAAATAGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1122
QY 1141 TCAGTTTCTGCTCTGCTGGGGGTATCTTAGTGGCTCTCATGAATCGGGTGGTGGTGGT 1200
DB 1123 TCAGTTTCTGCTCTGCTGGGGGTATCTTAGTGGCTCTCATGAATCGGGTGGTGGTGGT 1182
QY 1201 AATTTCTGCTGAGTTTCTGCTGGCAGTGGCGCTGGGACTTTGAGTGGTGGTGGTGGT 1260
DB 1183 AATTTCTGCTGAGTTTCTGCTGGCAGTGGCGCTGGGACTTTGAGTGGTGGTGGTGGT 1242
QY 1261 TACACCTTCTTCCACATCTCATGCAAGTCAACACATAGTCAAGCAGTGAAGAACAG 1320
DB 1243 TACACCTTCTTCCACATCTCATGCAAGTCAACACATAGTCAAGCAGTGAAGAACAG 1302
QY 1321 CAATGGAATGAAGAGAGACCTTTTCACTGATCTGCTCTCAAAACATAGAGAA 1380
DB 1303 CAATGGAATGAAGAGAGACCTTTTCACTGATCTGCTCTCAAAACATAGAGAA 1362
QY 1381 GTGCTTATTTGATTCACAGTGGAGGCTCTAACAGCTTAGGAGGCTGTGATTTATGT 1440
DB 1363 GTGCTTATTTGATTCACAGTGGAGGCTCTAACAGCTTAGGAGGCTGTGATTTATGT 1422
QY 1441 TTCTGTGAAATGCTTCTCAGATGATCAAAATTAAGATGAAGAAAGAAATC 1500
DB 1423 TTCTGTGAAATGCTTCTCAGATGATCAAAATTAAGATGAAGAAAGAAATC 1482
QY 1501 AGAAGAACTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
DB 1483 AGAAGAACTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1542
QY 1561 CTCACCTTTCAAAATGAGGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1620
DB 1543 CTCACCTTTCAAAATGAGGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 1602
QY 1621 GAGCAGACTCAGAGGCGCTCCACATTTGATTTCTCAGCAGCTGAGTCTGAGAGAG 1680
DB 1603 GAGCAGACTCAGAGGCGCTCCACATTTGATTTCTCAGCAGCTGAGTCTGAGAGAG 1662
QY 1681 AAGAGTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1663 AAGAGTCTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722
QY 1741 GGTCAAGAAATAATTCACATTTCCACGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1723 GGTCAAGAAATAATTCACATTTCCACGATGATGATGATGATGATGATGATGATGATGAT 1782
QY 1801 TTACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
DB 1783 TTACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1842


```
|||||
Db 781 TCCTCAGATTGATCAACAAATTTAAAGATAGAGAGAAAGAAATCAGAGAACCTGAAA 840
QY 1516 ATGATGATGATGAGATTAAGAGACAGTGTGCAAGATGATGAATCTCAACTTTCAACAA 1575
Db 841 ATGATGATGATGAGATTAAGAGACAGTGTGCAAGATGATGAATCTCAACTTTCAACAA 900
QY 1576 ATGAGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1635
Db 901 ATGAGAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 1636 AGCCCTCCCACTTTGATTTTCAAGACAGCTGCAAGTGTGGAAGAGAGAGGTCATGATAG 1695
Db 961 AGCCCTCCCACTTTGATTTTCAAGACAGCTGCAAGTGTGGAAGAGAGAGGTCATGATAG 1020
QY 1696 CTCATGCTCATCCACAGAGATGATCAATGAATATGATGATGATGATGATGATGATGATG 1755
Db 1021 CTCATGCTCATCCACAGAGATGATCAATGAATATGATGATGATGATGATGATGATGATG 1080
QY 1756 GCATTCACATTCACAGATACACTCGCCAGTCAGAGATCTCATTCACCCACATCATG 1815
Db 1081 GCATTCACATTCACAGATACACTCGCCAGTCAGAGATCTCATTCACCCACATCATG 1140
QY 1816 ACT 1818
Db 1141 ACT 1143

RESULT 6
US-10-198-846-10321
; Sequence 10321, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10321
; LENGTH: 1449
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1447, 1448, 1449
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10321
```

```
Query Match 26.0%; Score 900.4; DB 14; Length 1449;
Best Local Similarity 96.4%; Pred No. 9.9e-219;
Matches 985; Conservative 0; Mismatches 31; Indels 6; Gaps 6;

QY 1104 ATAGCCTGGTGTGGGTTTATAGCCATTCATCAGTTTCCTGCTGCTGGGG 1163
Db 199 ACAGCCTGGTGTGGGTTTATAGCCATTCATCAGTTTCCTGCTGCTGGGG 258
QY 1164 GTTATCTAGTCCCTCAGATGATCGGGTGTGTTTCAATTCCTCAGTTTCCTGTG 1223
Db 259 GTTATCTAGTCCCTCAGATGATCGGGTGTGTTTCAATTCCTCAGTTTCCTGTG 318
QY 1224 GCACTGGCGTGGGACTTTGAGTGGTGTGTTTACACCTTCCTCAGATTCAT 1283
Db 319 GCACTGGCGTGGGACTTTGAGTGGTGTGTTTACACCTTCCTCAGATTCAT 378
```

```
QY 1284 GCAAGTCCACCATAGTCTATAGCCATGAGAGAACCAAGCAATGGAATGAAAGAGAGACCA 1343
Db 379 GCAAGTCCACCATAGTCTATAGCCATGAGAGAACCAAGCAATGGAATGAAAGAGAGACCA 438
QY 1344 CTTTTCAGTCTATCTGTTCTCAAAACATAGAGAAAGTGCCTATTTTGTATTCACAGTGG 1403
Db 439 CTTTTCAGTCTATCTGTTCTCAAAACATAGAGAAAGTGCCTATTTTGTATTCACAGTGG 498
QY 1404 AAGGTCCTAACAGCTCTAGAGAGCGCTGATTTTCATGTTTCTTTGTAACATGCTCACA 1463
Db 499 AAGGTCCTAACAGCTCTAGAGAGCGCTGATTTTCATGTTTCTTTGTAACATGCTCACA 558
QY 1464 TTGATCAACAATTTAAAGATAGAGAAAGAAAGATCAGAGAAACCTGAAATGATGAT 1523
Db 559 TTGATCAACAATTTAAAGATAGAGAAAGAAAGATCAGAGAAACCTGAAATGATGAT 618
QY 1524 GATGCGAGATTAAGAGACAGTGTCCAAAGTATGAATCTCAACTTTTCAACAAATCAGGAG 1583
Db 619 GATGCGAGATTAAGAGACAGTGTCCAAAGTATGAATCTCAACTTTTCAACAAATCAGGAG 678
QY 1584 AAAGTAGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1643
Db 679 AAAGTAGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 738
QY 1644 CACTTTGATTTCTCAGCAGCGCTGCAGTCTTTGGAAGAGAGAGGTCATGATAGCTCATGCT 1703
Db 739 CACTTTGATTTCTCAGCAGCGCTGCAGTCTTTGGAAGAGAGAGGTCATGATAGCTCATGCT 798
QY 1704 CATCCACAGGAAGTCTACATGAATATGATGATGATGATGATGATGATGATGATGATGATG 1763
Db 799 CATCCACAGGAAGTCTACATGAATATGATGATGATGATGATGATGATGATGATGATGATG 858
QY 1764 CATTTCCAGATACACTCGCCAGTCAGACGATCTCATTCACCAACCATCATGATGATGAT 1823
Db 859 CATTTCCAGATACACTCGCCAGTCAGACGATCTCATTCACCAACCATCATGATGATGAT 918
QY 1824 CATATTTCCATCATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1882
Db 919 CATATTTCCATCATCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 978
QY 1883 CTCCTGGGAGGAGCTGAAAGATGCCGGCTGCGCACCTTTGGCTGGATGGTGAATGGG 1942
Db 979 CTCCTGGGAGGAGCTGAAAGATGCCGGCTGCGCACCTTTGGCTGGATGGTGAATGGG 1038
QY 1943 TGAATGCGCTGCACAA-TTTCAGCGATGCGCTAGCAATGG-TGCTGCTTTTACTGAA-GG 1999
Db 1039 TGAATGCGCTGCACAA-TTTCAGCGATGCGCTAGCAATGG-TGCTGCTTTTACTGAA-GG 1098
QY 2000 CTTATCAAGTGGTTTAAAGTACTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
Db 1099 CTTATCAAGTGGTTTAAAGTACTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157
QY 2060 AGGTGACTTTGCTGTTTCTACTAAAGCTGGCATGACCGTTAAGCAGGCTGCTCTTATAA 2119
Db 1158 AGGAGACTAAAGTCTGTCTACTAACAAGGTTGGCATTAACG-TAAGCGCGCTGCTCTCTTCA 1216
QY 2120 TG 2121
Db 1217 TG 1218

RESULT 7
US-10-198-846-13805
; Sequence 13805, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
```


Db	723	AAGTGGAAATCCCTTCCAAAGAGCTAATTCATTTCACAAATAGGCTGAGTTGGGGTATATA	782
QY	1128	GCACATTTCCATCATCAGTTTCTTCGTCTCTGCTGGGGGTATCTTAGTGCTCTCATGAAT	1187
Db	793	GGCAATTTCCATCAATCAGTACTTG--CACTGCGGGGTAACTAATACGGCCCTTAATGAT	840
QY	1188	CGGGTGTTTTCAAATTTCTCCTGAGTTTCTTGTTGGCACTGCGCGGTG	1236
Db	841	CGGGGTTCACAAATTCCTGGCAATTCCTTGACTGCGGTGGCACTTG	889
RESULT 8			
US-10-198-846-8490			
; Sequence 8490, Application US/10198846			
; Publication No. US20030099974A1			
; GENERAL INFORMATION:			
; APPLICANT: Lillie, James			
; APPLICANT: Xu, Yongyao			
; APPLICANT: Wang, Youzhen			
; APPLICANT: Steinmann, Kathleen			
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS			
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND			
; TITLE OF INVENTION: THERAPY OF BREAST CANCER			
; FILE REFERENCE: MRI-049			
; CURRENT APPLICATION NUMBER: US/10/198,846			
; CURRENT FILING DATE: 2002-07-18			
; PRIOR APPLICATION NUMBER: 60/306,220			
; PRIOR FILING DATE: 2001-07-18			
; NUMBER OF SEQ ID NOS: 14084			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 8490			
; LENGTH: 606			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: 1, 44, 516, 571, 600, 606			
; OTHER INFORMATION: n = A,T,C or G			
US-10-198-846-8490			
Query Match			
Best Local Similarity 16.7%; Score 578.4; DB 14; Length 606;			
Matches 590; Conservative 0; Mismatches 5; Indels 1; Gaps 1;			
QY	1179	CTCATGAATCGGGTGTTTTCAAATTTCTCCTGAGTTTCTTGTTGGCACTGCGCGTTGGG	1238
Db	11	CGCATGAATCGGGTGTTTTCAAATTTCTCCTG--NPTTCCTTGTTGGCACTGCGCGTTGGG	69
QY	1239	ACTTTGATGGTGATGCGTTTTTTTACACCTTCTTCACATTCATCATGCAAGTCACACCAT	1298
Db	70	ACTTTGATGGTGATGCGTTTTTTTACACCTTCTTCACATTCATCATGCAAGTCACACCAT	129
QY	1299	AGTCATAGCCATCAGAACACCGCAATGGAAATGAAAGAGAGACCATTTTCAGTCATCTG	1358
Db	130	AGTCATAGCCATCAGAACACCGCAATGGAAATGAAAGAGAGACCATTTTCAGTCATCTG	189
QY	1359	TCATTCTCAAAACATAGAAGAAGTGCCATTATTTGATTTCCACGTGGAAAGGCGTTAAACGCT	1418
Db	190	TCATTCTCAAAACATAGAAGAAGTGCCATTATTTGATTTCCACGTGGAAAGGCGTTAAACGCT	249
QY	1419	CTAGAGGCGCTGATTTTCATGTTTTCCTTTGTAACATGTCCTCACAATGATGATGATGAGATTAA	1478
Db	250	CTAGAGGCGCTGATTTTCATGTTTTCCTTTGTAACATGTCCTCACAATGATGATGATGAGATTAA	309
QY	1479	AAACATAGAAGAAAAGAAAGAAATCAGAGAAACCTGAAATGATGATGATGAGATTAA	1538
Db	310	AAAGATAGAAGAAAAGAAAGAAATCAGAGAAACCTGAAATGATGATGATGAGATTAA	369
QY	1539	AAGCAGTTGCCAAGTATGTAATCTCAACTTTCAACAAATGAGGAGAAAGTAGATACAGAT	1598
Db	370	AAGCAGTTGCCAAGTATGTAATCTCAACTTTCAACAAATGAGGAGAAAGTAGATACAGAT	429
QY	1599	GATCGAACTGAAGGCTATTTACGAGACGAGACTCACAAAGACCGCTCCACACTTTGATTTCTCAG	1658


```
Db 430 GATCGACTGAGGCTATTACAGCAGACTCACAGAGCCCTCCACCTTGAATTCTCAG 489
QY 1659 CAGCCCTGCAGCTTGGGAAGAGAGGTCATGATAGTCATGTCATCCACAGGAATC 1718
Db 490 CAGCCTGCAGCTTGGGAAGAGAGGTCATGATAGTCATGTCATCCACAGGAATC 549
QY 1719 TACAATGAATATGTACCACAGAGGTCGAAGATAAATGCCATTCACATTTCCAGA 1774
Db 550 TACAATGAATATGTACCACAGAGGTCGAAGATAAATGCCATTCACATTTCCAGA 605

RESULT 9
US-10-198-846-11215
; Sequence 11215, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR APPLICATION NUMBER: 2002-07-18
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11215
; LENGTH: 1598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1594, 1595, 1596, 1597, 1598
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11215

Query Match 16.7%; Score 578.4; DB 14; Length 1598;
Best Local Similarity 91.3%; Pred. No. 1.5e-136;
Matches 725; Conservative 0; Mismatches 56; Indels 13; Gaps 10;
QY 2245 TTGATATGCTACTGAATGCTGCACAAATGATGCTAGTGACCATGGATGGATGGCGCTGGG 2304
Db 682 TTTTAAAGGTACTCGAAATGCTGCACAAATGATGCTAGTGACCATGGATGGATGGCGCTGGG 741
QY 2305 GGTATTTCCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATTAATGTTACTTATT 2364
Db 742 GGTATTTCCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATTAATGTTACTTATT 799
QY 2365 CCATATTGAACATAAAATCGTGTTCGTATATAAATTCATGTTAAGTTTAAATGCTAGA 2424
Db 800 CCATATTGAACATAAAATCGTGTTCGTATATAAATTCATGTTAAGTTTAAATGCTAGA 859
QY 2425 GTAGCTTAAAAGTTGTCATAGTTTCAGTAGGTCATAGGAGATGAGTTTGTATGCTGTA 2484
Db 860 GTAGCTTAAAAGTTGTCATAGTTTCAGTAGGTCATAGGAGATGAG-TTGTATGCTGTA 918
QY 2485 CTATGCAGCGTTTAAAAGTTAGTGGTTCGTGATTTTTCATATGAATATGCGTCTGTT 2544
Db 919 CTATGCAGCGTTTAAAAGTTAGTGGTTCGTGATTTTTCATATGAATATGCGTCTGTT 976
QY 2545 ACAAGTCAGTTTAAAGGTACGTTTAAATTAATTAAGTTAATTCATCTTTGGAGATAAAATCT 2604
Db 977 ACAAGTCAGTTTAAAGGTACGTTTAAATTAATTAAGTTAATTCATCTTTGGAGATAAAATCT 1036
QY 2605 GTATGTCATATCCCGGTATATACAGTTTATATGTAATAACAGAGATTTGGCATGACAT 2664
Db 1037 GTATGTCATATCCCGGTATATACAGTTTATATGTAATAACAGAGATTTGG-ATGACAT 1095
```

```
QY 2665 GTTCTGTATGTTTTCAGGAAAAATGCTTTAATGCTTTTCAAGAACTAAACACAGTTATT 2724
Db 1096 GTTCTGTATGTTTTCAGGAAAAATGCTTTAATGCTTTTCAAGAACTAAACACAGTTATT 1155
QY 2725 CTTATACTGGATTTTAGTGTCTCTGAAGAACTGCTCGTGTCTTTAGGAATTAAGAATGTGCATG 2784
Db 1156 CTTATACTGGATTTTAGTGTCTCTGAAGAACTG-TGGTGTCTTAGGAATAAGAATGTGCATG 1214
QY 2785 AGCCTAAATATCCACAAAGAACTTATACCTGAATTTAAGCAAGAAATAAAGAGAGAAAGA 2844
Db 1215 AGCCTAAATATCCACAAAGAACTTATACCTGAATTTAAGCAAGAAATAAAGAGAGAAAGA 1274
QY 2845 GAAGATCTGAGAAATTGGGGAGGATAGATCTTTATAAAATCACAAAATTTGTTGTAAA 2904
Db 1275 GAAGATCTGAGAA-TGGGGAGGATAGA-TCTTTAATAAATCACAAATTCGT--GTAA 1330
QY 2905 TTAGAGGGGAGAAATTTAGAAATTAAGTA-TAAAAAGGCAGAAATTTAGTATAGAGTACATC 2963
Db 1331 TTAGAGGGGAGAAATTCAGAACTTACGTAATTAACAAGGGCGAAATTTAGTCTTTGAAGTCCATC 1390
QY 2964 ATTAACATTTTGTCCAGGATTTATTTCCCGTAAAAACGTAGTGAGCACTCTCATATACTA 3023
Db 1391 CTCACACCTCTCTGTC-GGATTTATTTCCGTTAAAAACGTTGTGGCCCTTTCTTAATCAATT 1449
QY 3024 ATTAGTGTACATTT 3037
Db 1450 TCGTGTCTTACTT 1463

RESULT 10
US-09-867-701-2021
; Sequence 2021, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2021
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(483)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-2021

Query Match 11.7%; Score 404.8; DB 10; Length 483;
Best Local Similarity 92.1%; Pred. No. 1.2e-92;
Matches 444; Conservative 0; Mismatches 35; Indels 3; Gaps 2;
QY 2541 TGTTCACAAAGTCAGTTAAAGTCAGTTTAAATTAATTAAGTTTAAATTAATTAAGTTTAA 2600
Db 2 TGTTCACAAAGTCAGTTAAAGTCAGTTTAAATTAATTAAGTTTAAATTAATTAAGTTTAA 61
QY 2601 ATCTGTATGTCGAATTCACCGGTATTAACAGTTTATTAATGTAATAAACAAGAGATTTGGCATG 2660
Db 62 ATCTGTATGTCGAATTCACCGGTATTAACAGTTTATTAATGTAATAAACAAGAGATTTGGCATG 121
QY 2661 ACATGTTCTGTATGTTTTCAGGAAAAATGCTTTAATGCTTTTCAAGAACTAACACAGT 2720
Db 122 ACATGTTCTGTATGTTTTCAGGAAAAATGCTTTAATGCTTTTCAAGAACTAACACAGT 181
QY 2721 TATTCCTATACCTGGATTTTAGTCTCTGGAAGAACTGCTGGTGTGTTTAGAATAAAGATG 2780
Db 182 TATTCCTATACCTGGATTTTAGTCTCTGGAAGAACTGCTGGTGTGTTTAGAATAAAGATG 241
```

QY 2781 CATGAGCCTAAATACCAAGAGCTTATCTACTGAATTTAGCAAGAAATTAAGGAGAA 2840
 |||||
 Db 242 CATGAGCCTAAATACCAAGAGCTTATCTACTGAATTTAGCAAGAAATTAAGGAGAA 301
 |||||
 QY 2841 AAGAGAAGATC--TGAGAATTGGGAGGCATAGATCTCTATAAAATACAAAATTTCT 2898
 |||||
 Db 302 AAGAGAAGATCCTGAGAATTGGGAGGCATAGATCTCTATAAAATACAAAATTTCT 361
 |||||
 QY 2899 TGTAAA--TAGAGGGGAGAAATTTAGAAATTAAGATATAAAAGCGCAGATTTAGTAGT 2957
 |||||
 Db 362 TGTAAATTTAGAGGGGAGAAATTTAGAAATTAAGATATAAAAGCGCAGATTTAGTAGT 421
 |||||
 QY 2958 ACATTATTAACATTTTGTGAGATTTTCCCGTAAACGCTAGTGAGCACTCTCAT 3017
 |||||
 Db 422 ACNTCCNTTACATTTTGGCGGATTTTCCCGTAAACGCTAGTGAGCACTCTCAT 481
 |||||
 QY 3018 AT 3019
 ||
 Db 482 AT 483
 ||
 RESULT 11
 US-09-778-320-94/c
 ; Sequence 94, Application US/09778320
 ; Patent No. US20010034052A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; CURRENT FILING DATE: 2001-02-06
 ; CURRENT APPLICATION NUMBER: US/09/778,320
 ; NUMBER OF SEQ ID NOS: 301
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 94
 ; TYPE: DNA
 ; LENGTH: 395
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(395)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-778-320-94
 Query Match 11.3%; Score 390.8; DB 9; Length 395;
 Best Local Similarity 99.2%; Pred. No. 3.8e-59;
 Matches 392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2563 ACCTTTTAATTAAGTTATCTCTATCTGGAGATAAAATCTGTATGTGCAATTCACCGG 2622
 |||||
 Db 395 ACCTTTTAATTAAGTTATCTCTATCTGGAGATAAAATCTGTATGTGCAATTCACCGG 336
 |||||
 QY 2623 TATTACAGTTTATTATGTAACAGAGATTGGCATGACATGTTCTGTATGTTCAGGG 2682
 |||||
 Db 335 TATTACAGTTTATTATGTAACAGAGATTGGCATGACATGTTCTGTATGTTCAGGG 276
 |||||
 QY 2683 AAAAATGCTTTTAATGCTTTTCAAGAACTAACACAGTTATTCCTTACTGGAATTTAGG 2742
 |||||
 Db 275 AAAAATGCTTTTAATGCTTTTCAAGAACTAACACAGTTATTCCTTACTGGAATTTAGG 216
 |||||
 QY 2743 TCTCTGAAGAACTGCTGTGTGTAGGAATAAGAAATGTCATGAAGCCTAAATACCAAGA 2802
 |||||
 Db 215 TCTCTGAAGAACTGCTGTGTGTAGGAATAAGAAATGTCATGAAGCCTAAATACCAAGA 156
 |||||
 QY 2803 AAGCTTACTGAATTTAGCAAGAAATTAAGAGAGAAAGAGAAATCTGGAATTTGG 2862
 |||||
 Db 155 AAGCTTACTGAATTTAGCAAGAAATTAAGAGAGAAAGAGAAATCTGGAATTTGG 96
 |||||
 QY 2863 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 2922
 |||||
 Db 95 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 36
 |||||
 QY 2923 GAATTAAGTATAAAAGGAGCAATTTAGTAGAGT 2957
 |||||
 Db 35 GAATTAAGTATAAAAGGAGCAATTTANATAGAGT 1
 |||||

Db 155 AAGCTTACTGAATTTAGCAAGAAATTAAGAGAGAAAGAGAAATCTGGAATTTGG 96
 |||||
 QY 2863 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 2922
 |||||
 Db 95 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 36
 |||||
 QY 2923 GAATTAAGTATAAAAGGAGCAATTTAGTAGAGT 2957
 |||||
 Db 35 GAATTAAGTATAAAAGGAGCAATTTANATAGAGT 1
 |||||
 RESULT 12
 US-09-910-689-94/c
 ; Sequence 94, Application US/09910689
 ; Patent No. US20020081609A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Jiang, Yuqiu
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: Mitcham, Jennifer
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Harlocker, Susan L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; FILE OF INVENTION: DIAGNOSIS OF BREAST CANCER
 ; CURRENT FILING DATE: 2001-07-20
 ; CURRENT APPLICATION NUMBER: US/09/910,689
 ; NUMBER OF SEQ ID NOS: 307
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 94
 ; LENGTH: 395
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 9
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-910-689-94
 Query Match 11.3%; Score 390.8; DB 9; Length 395;
 Best Local Similarity 99.2%; Pred. No. 3.8e-59;
 Matches 392; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2563 ACCTTTTAATTAAGTTATCTCTATCTGGAGATAAAATCTGTATGTGCAATTCACCGG 2622
 |||||
 Db 395 ACCTTTTAATTAAGTTATCTCTATCTGGAGATAAAATCTGTATGTGCAATTCACCGG 336
 |||||
 QY 2623 TATTACAGTTTATTATGTAACAGAGATTGGCATGACATGTTCTGTATGTTCAGGG 2682
 |||||
 Db 335 TATTACAGTTTATTATGTAACAGAGATTGGCATGACATGTTCTGTATGTTCAGGG 276
 |||||
 QY 2683 AAAAATGCTTTTAAGTCTTTTCAAGAACTAACACAGTTATTCCTTACTGGAATTTAGG 2742
 |||||
 Db 275 AAAAATGCTTTTAAGTCTTTTCAAGAACTAACACAGTTATTCCTTACTGGAATTTAGG 216
 |||||
 QY 2743 TCTCTGAAGAACTGCTGTGTGTAGGAATAAGAAATGTCATGAAGCCTAAATACCAAGA 2802
 |||||
 Db 215 TCTCTGAAGAACTGCTGTGTGTAGGAATAAGAAATGTCATGAAGCCTAAATACCAAGA 156
 |||||
 QY 2803 AAGCTTACTGAATTTAGCAAGAAATTAAGAGAGAAAGAGAAATCTGGAATTTGG 2862
 |||||
 Db 155 AAGCTTACTGAATTTAGCAAGAAATTAAGAGAGAAAGAGAAATCTGGAATTTGG 96
 |||||
 QY 2863 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 2922
 |||||
 Db 95 GGAGGCATAGTTCTTATATAAATCACAATAATTTGTTTAAATTAGAGGGGAGAAATTTA 36
 |||||
 QY 2923 GAATTAAGTATAAAAGGAGCAATTTAGTAGAGT 2957
 |||||
 Db 35 GAATTAAGTATAAAAGGAGCAATTTANATAGAGT 1
 |||||

US-10-010

RESULT 14

```

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22272
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(504)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-22272

Query Match 11.3%; Score 399.6; DB 11; Length 50
Best Local Similarity 96.8%; Pred. No. 9.3e-89;
Matches 451; Conservative 0; Mismatches 9; Indels 0

QY 2938 AGGCAGAAATTAGTATAGAGTACATTCATTAAACATTTTGTTCAGGAGTATT
DB 39 AGGCAGAAATTAGTATAGAGTACATTCATTAAACATTTTGTTCAGGAGTATT
QY 2998 AACGTAGTGGACAC-TCTCATATACIAAATT--AGTGACATTTTAACTTTGT
DB 99 AACGTAGTGGACACTTTTCATATACTAATTTAGTTGTACATTTTAACTTTGT
QY 3055 AAATCTAAATATATTAAATGAATTCAGCAATATA-CACTTGACCAGAAA
DB 159 AAATCTAAATATATTTAAGATTTCAGCAATATATCACTTGACCAGAAA
QY 3114 CAAAATGTCGTGCGGGTTATATACCAGATGAGTACAGTAGAGTAG-TTTAT
DB 219 CAAAATGTCGTGCGGG-TATATACCAGATGAGTACAGTAGAGTAGTTTAT
QY 3173 GACTGGGTTATTCGCAAGTTATATATACCAAAAGCTGTATGACTGGAGTT
DB 278 GACTGGGTTATTCGCAAGTTATATATACCAAAAGCTGTATGACTGGAGTT
QY 3233 CTGGTTTACAAAATTATCAGAGTAGTAAACCTTTGATATATATGAGGATAT
DB 338 CTGGTTTACAAATATTACAGAGTAGTACAACTTTGATATATATGAGGATAT
QY 3293 ACTAAGTATCATTTGATTCGATTCAGAAAGTAACTTTGATATCTCTCAGTGC
DB 398 ACTAAGTATCATTTGATTCGATTCATTAAGTAACTGTGATATCTCTCAGTGC
QY 3353 ATCATTTGAGCAATCTCTTTTATATACGGTACTGTAGCCATACCTA 3398
DB 458 ATCATTTGAGCAATGGCTTTTATATACGGTACTGTAGCCATACCTA 503

RESULT 15
US-09-827-669-78
; Sequence 78, Application US/09827669
; Patent No. US20020009739A1
; GENERAL INFORMATION:
; APPLICANT: Gleese, Klaus
; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER
; FILE REFERENCE: 200130.460
; CURRENT APPLICATION NUMBER: US/09/827,669
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 09/417,615
; PRIOR APPLICATION NUMBER: 1999-10-13
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 78

```

```
; LENGTH: 385
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-827-669-78

Query Match      10.4%; Score 360; DB 9; Length 385;
Best Local Similarity 98.2%; Pred. No. 2.7e-81;
Matches 375; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2453 TAGTCATAGGAGATGAGTTTGTCTACTATGACAGCGTTTAAAGTTAGTGGGTTT 2512
Db 1 TCGGTCATAGGAGAG--GATTTGTATGCTGCTACTATGACAGCGTTTAAAGTTAGTGGGTTT 58

QY 2513 TGTGATTTTGTATTGATATATTGCTGCTGTTACAAAGTCAGTTAAAGGTACGTTTAAAT 2572
Db 59 TGTGATTTTGTATTGATATATTGCTGCTGTTACAAAGTCAGTTAAAGGTACGTTTAAAT 118

QY 2573 ATTTAAGTTATTCTATCTTTGGAGATAAAATCTGTATGTCCAATTCACCGGTATTACCAAGT 2632
Db 119 ATTTAAGTTATTCTATCTTTGGAGATAAAATCTGTATGTCCAATTCACCGGTATTACCAAGT 178

QY 2633 TTATATGCTAAACAAGACAGATTGGCATGCACATGCTGCTATCTTTACGGGAAAAATGCTCT 2692
Db 179 TTATATGCTAAACAAGACAGATTGGCATGCACATGCTGCTATCTTTACGGGAAAAATGCTCT 238

QY 2693 TTAATGCTTTTTCAGAACTAACACAGTTATTCTCTACTGATTTAGGTCCTCTGAAGA 2752
Db 239 TTAATGCTTTTTCAGAACTAACACAGTTATTCTCTACTGATTTAGGTCCTCTGAAGA 298

QY 2753 ACTGCTGGTGTATTAGGATATAGAAATGTGCATGAGCCTAAATACCAAGAAAGCTTATAC 2812
Db 299 ACTGCTGGTGTATTAGGATATAGAAATGTGCATGAGCCTAAATACCAAGAAAGCTTATAC 358

QY 2813 TGAATTTAAGCAAGAAATATA 2834
Db 359 TGAATTTAAGCAAGAAATATA 380
```

Search completed: September 12, 2003, 22:12:41
Job time : 804.001 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 12, 2003, 16:38:39 ; Search time 126.286 Seconds
(without alignments)
7926.909 Million cell updates/sec

Title: US-09-642-034-4
Perfect score: 2268
Sequence: 1 atgcgagggaagtatctgt.....tggttgataaaattctgt 2268

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1263	55.7	1310	1 US-08-311-023-1	Sequence 1, Appli
2	1263	55.7	2404	1 US-08-311-023-3	Sequence 3, Appli
3	110.6	4.9	1888	4 US-09-599-360B-38	Sequence 38, Appli
c 4	59.8	2.6	285	3 US-08-990-571-64	Sequence 64, Appli
c 5	59.8	2.6	285	4 US-09-528-784A-64	Sequence 64, Appli
c 6	59.8	2.6	285	4 US-09-569-098A-64	Sequence 64, Appli
c 7	59.8	2.6	342	3 US-08-990-571-65	Sequence 65, Appli
c 8	59.8	2.6	342	4 US-09-528-784A-65	Sequence 65, Appli
c 9	59.8	2.6	342	4 US-09-569-098A-65	Sequence 65, Appli
c 10	59.8	2.6	351	3 US-08-990-571-60	Sequence 60, Appli
c 11	59.8	2.6	351	4 US-09-528-784A-60	Sequence 60, Appli
c 12	59.8	2.6	356	3 US-09-569-098A-60	Sequence 60, Appli
c 13	59.8	2.6	356	3 US-08-990-571-63	Sequence 63, Appli
c 14	59.8	2.6	356	4 US-09-528-784A-63	Sequence 63, Appli
c 15	59.8	2.6	356	4 US-09-569-098A-63	Sequence 63, Appli
c 16	59.8	2.6	358	3 US-08-990-571-58	Sequence 58, Appli
c 17	59.8	2.6	358	4 US-09-528-784A-58	Sequence 58, Appli
c 18	59.8	2.6	358	4 US-09-569-098A-58	Sequence 58, Appli
c 19	59.8	2.6	363	3 US-08-990-571-66	Sequence 66, Appli
c 20	59.8	2.6	363	3 US-08-990-571-67	Sequence 67, Appli
c 21	59.8	2.6	363	4 US-09-528-784A-66	Sequence 66, Appli
c 22	59.8	2.6	363	4 US-09-528-784A-67	Sequence 67, Appli
c 23	59.8	2.6	363	4 US-09-569-098A-66	Sequence 66, Appli
c 24	59.8	2.6	363	4 US-09-569-098A-67	Sequence 67, Appli
c 25	59.8	2.6	409	3 US-08-990-571-59	Sequence 59, Appli
c 26	59.8	2.6	409	4 US-09-528-784A-59	Sequence 59, Appli
c 27	59.8	2.6	409	4 US-09-569-098A-59	Sequence 59, Appli

c 28	59.8	2.6	410	3 US-08-990-571-61	Sequence 61, Appli
c 29	59.8	2.6	410	4 US-09-528-784A-61	Sequence 61, Appli
c 30	59.8	2.6	410	4 US-09-569-098A-61	Sequence 61, Appli
c 31	59.8	2.6	416	3 US-08-990-571-62	Sequence 62, Appli
c 32	59.8	2.6	416	4 US-09-528-784A-62	Sequence 62, Appli
c 33	59.8	2.6	416	4 US-09-569-098A-62	Sequence 62, Appli
c 34	59.8	2.6	1271	3 US-08-845-258-5	Sequence 5, Appli
c 35	59.8	2.6	1271	3 US-08-845-258-42	Sequence 42, Appli
c 36	59.8	2.6	1271	3 US-08-990-571-5	Sequence 5, Appli
c 37	59.8	2.6	1271	3 US-08-990-571-42	Sequence 42, Appli
c 38	59.8	2.6	1271	4 US-08-723-142A-5	Sequence 5, Appli
c 39	59.8	2.6	1271	4 US-08-723-142A-42	Sequence 42, Appli
c 40	59.8	2.6	1271	4 US-09-528-784A-5	Sequence 5, Appli
c 41	59.8	2.6	1271	4 US-09-528-784A-42	Sequence 42, Appli
c 42	59.8	2.6	1271	4 US-09-569-098A-5	Sequence 5, Appli
c 43	59.8	2.6	1271	4 US-09-569-098A-42	Sequence 42, Appli
c 44	59.8	2.6	1821	3 US-08-845-258-6	Sequence 6, Appli
c 45	59.8	2.6	1821	3 US-08-990-571-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-311-023-1
; Sequence 1, Application US/083111023
; Patent No. 5693465
; GENERAL INFORMATION:
; APPLICANT: MANNING, David Lockwood
; APPLICANT: NICHOLSON, Robert Ian
; APPLICANT: GEE, Julia Margaret
; APPLICANT: GREEN, Christopher Douglas
; TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOUR OF
; TITLE OF INVENTION: BREAST TUMOURS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Young & Thompson
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,023
; FILING DATE: 22-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: Reg. No. 5693465 32.925
; REFERENCE/DOCKET NUMBER: WCM.56
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1310 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1293
US-08-311-023-1

Query Match 55.7%; Score 1263; DB 1; Length 1310;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1285; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

OY 2131 C 2131
Db 1632 C 1632

RESULT 4

US-08-990-571-64/c
; Sequence 64, Application US/08990571
; Patent No. 6214971
; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G. et al.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B. M
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSES: SED AND BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,571
; FILING DATE: 11-DEC-1997
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.426C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 64:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-990-571-64

Query Match 2.6%; Score 59.8; DB 3; Length 285;
Best Local Similarity 65.2%; Pred. No. 3.2e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 268 ATCCATATACACCATGAGCATCTCAGACCATCATCTCAGACCATCATCTCAGACCATGAG 327
Db 192 AACCAAGTTCCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATG 133
OY 328 CGTCACTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGAC 387
Db 132 CTTCACCTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCA 73
OY 388 TCTCACCATATCAT 402
Db 72 TAGGCCACCATGCTT 58

RESULT 5

US-08-528-784A-64/c
; Sequence 64, Application US/09528784A
; Patent No. 6451315
; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS

; TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
; FILE REFERENCE: 210121.426C4
; CURRENT APPLICATION NUMBER: US/09/528,784A
; CURRENT FILING DATE: 2000-03-17
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Babesia microti
US-09-528-784A-64

Query Match 2.6%; Score 59.8; DB 4; Length 285;
Best Local Similarity 65.2%; Pred. No. 3.2e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 268 ATCCATATACACCATGAGCATCTCAGACCATCATCTCAGACCATGAGCATCTCAGACCATGAG 327
Db 192 AACCAAGTTCCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATG 133
OY 328 CGTCACTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGACCATG 387
Db 132 CTTCACCTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCA 73
OY 388 TCTCACCATATCAT 402
Db 72 TAGGCCACCATGCTT 58

RESULT 6

US-09-569-098A-64/c
; Sequence 64, Application US/09569098A
; Patent No. 6569433
; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Sleath, Paul R.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Homer, Mary
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
; FILE REFERENCE: 210121.426C5
; CURRENT APPLICATION NUMBER: US/09/569,098A
; CURRENT FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 285
; TYPE: DNA
; ORGANISM: Babesia microti
US-09-569-098A-64

Query Match 2.6%; Score 59.8; DB 4; Length 285;
Best Local Similarity 65.2%; Pred. No. 3.2e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

OY 268 ATCCATATACACCATGAGCATCTCAGACCATCATCTCAGACCATGAGCATCTCAGACCATGAG 327
Db 192 AACCAAGTTCCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATG 133
OY 328 CGTCACTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGACCATGAGCATCTCAGACCATG 387
Db 132 CTTCACCTAGGCCACCATGCTTCACTAGGCCACCATGCTTCACTAGGCCACCATGCTTCA 73
OY 388 TCTCACCATATCAT 402
Db 72 TAGGCCACCATGCTT 58

RESULT 7

US-08-990-571-65/c

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,571
FILING DATE: 11-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 351 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-990-571-60

Query Match 2.6%; Score 59.8; DB 3; Length 351;
Best Local Similarity 65.2%; Pred. No. 3.6e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 268 ATCCATATACACCATGACACGACGATCATCTCAGACGACGATCATCTCAGACCATGAG 327
DB 210 AACAGCTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACG 151
QY 328 CGTCACTCAGACCATGACATCATCTCAGACGACGATCATCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACGATTCAC 91
QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACACGCTT 76

RESULT 11
US-09-528-784A-60/c
Sequence 60, Application US/09528784A
Patent No. 6451315
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
AND TREATMENT OF B. MICROTI INFECTION
FILE REFERENCE: 210121.426C4
CURRENT APPLICATION NUMBER: US/09/528,784A
CURRENT FILING DATE: 2000-03-17
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 60
LENGTH: 351
TYPE: DNA
ORGANISM: Babesia microti
US-09-528-784A-60

Query Match 2.6%; Score 59.8; DB 4; Length 351;
Best Local Similarity 65.2%; Pred. No. 3.6e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 268 ATCCATATACACCATGACACGACGATCATCTCAGACGACGATCATCTCAGACCATGAG 327
DB 210 AACAGCTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACG 151
QY 328 CGTCACTCAGACCATGACATCATCTCAGACGACGATCATCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACGATTCAC 91
QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACACGCTT 76

RESULT 12
US-09-569-098A-60/c
Sequence 60, Application US/09569098A
Patent No. 6569433
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: McNeill, Patricia D.
APPLICANT: Homer, Mary
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
AND TREATMENT OF B. MICROTI INFECTION
FILE REFERENCE: 210121.426C5
CURRENT APPLICATION NUMBER: US/09/569,098A
CURRENT FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 110
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 60
LENGTH: 351
TYPE: DNA
ORGANISM: Babesia microti
US-09-569-098A-60

Query Match 2.6%; Score 59.8; DB 4; Length 351;
Best Local Similarity 65.2%; Pred. No. 3.6e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 268 ATCCATATACACCATGACACGACGATCATCTCAGACGACGATCATCTCAGACCATGAG 327
DB 210 AACAGCTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACG 151
QY 328 CGTCACTCAGACCATGACATCATCTCAGACGACGATCATCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACACGATTCATCTAGGCCACACGATTCATCTAGGCCACACGATTCAC 91
QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACACGCTT 76

RESULT 13
US-08-990-571-63/c
Sequence 63, Application US/08990571
Patent No. 6214371
GENERAL INFORMATION:
APPLICANT: Reed, Steven G. et al.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMENT OF B
NUMBER OF SEQUENCES: 79
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/990,571
FILING DATE: 11-DEC-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.426C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-990-571-63

Query Match 2.6%; Score 59.8; DB 3; Length 356;
Best Local Similarity 65.2%; Pred. No. 3.7e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 268 ATCCATATACACCATGACGACCATCTCAGACGACGATCTCAGACCATGAG 327
DB 210 AACCATGTCACAGGCGCCACGCTTCTACTAGGCCACCGAGCTTCTAGGCCACCGAG 151
QY 328 CGTCACTCAGACCATGACGACCATCTCAGACGACGATCTCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTAC 91
QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACCGAGCTT 76

RESULT 14
US-09-528-784A-63/C
Sequence 63, Application US/09528784A
Patent No. 6451315
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: McNeill, Patricia D.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
FILE REFERENCE: 210121.426C4
CURRENT APPLICATION NUMBER: US/09/528,784A
CURRENT FILING DATE: 2000-03-17
NUMBER OF SEQ ID NOS: 90
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 63
LENGTH: 356
TYPE: DNA
ORGANISM: Babesia microti
US-09-528-784A-63

Query Match 2.6%; Score 59.8; DB 4; Length 356;
Best Local Similarity 65.2%; Pred. No. 3.7e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 268 ATCCATATACACCATGACGACCATCTCAGACGACGATCTCAGACCATGAG 327
DB 210 AACCATGTCACAGGCGCCACGCTTCTACTAGGCCACCGAGCTTCTACTAGGCCACCGAG 151
QY 328 CGTCACTCAGACCATGACGACCATCTCAGACGACGATCTCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTAC 91

QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACCGAGCTT 76

RESULT 15

US-09-569-098A-63/C
Sequence 63, Application US/09569098A
Patent No. 6569433
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond L.
APPLICANT: Sleath, Paul R.
APPLICANT: McNeill, Patricia D.
APPLICANT: Homer, Mary
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
TITLE OF INVENTION: AND TREATMENT OF B. MICROTI INFECTION
FILE REFERENCE: 210121.426C5
CURRENT APPLICATION NUMBER: US/09/569,098A
CURRENT FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 110
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 63
LENGTH: 356
TYPE: DNA
ORGANISM: Babesia microti
US-09-569-098A-63

Query Match 2.6%; Score 59.8; DB 4; Length 356;
Best Local Similarity 65.2%; Pred. No. 3.7e-07;
Matches 88; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 268 ATCCATATACACCATGACGACCATCTCAGACGACGATCTCAGACCATGAG 327
DB 210 AACCATGTCACAGGCGCCACGCTTCTACTAGGCCACCGAGCTTCTACTAGGCCACCGAG 151
QY 328 CGTCACTCAGACCATGACGACCATCTCAGACGACGATCTCTGACCATGATCATCAC 387
DB 150 CTTCACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTACTAGGCCACCGAGCTTCTAC 91
QY 388 TCTCACCATAATCAT 402
DB 90 TAGGCCACCGAGCTT 76

Search completed: September 12, 2003, 16:55:42
Job time : 128.286 secs

C	17	240.8	10.6	469	14	US-10-198-846-7048	Sequence 408, App
	18	235	10.4	635	14	US-10-198-846-7649	Sequence 7649, App
	19	146.8	6.5	479	14	US-10-198-846-1957	Sequence 1957, App
	20	139	6.1	1598	14	US-10-198-846-11215	Sequence 11215, App
C	21	132	5.8	550	14	US-10-198-846-5880	Sequence 5880, App
	22	128.4	5.7	751	14	US-10-198-846-4814	Sequence 4814, App
	23	127.6	5.6	4573	12	US-09-880-107-1649	Sequence 1649, App
	24	127.6	5.6	4573	12	US-09-880-107-1649	Sequence 1649, App
	25	126	5.6	4698	12	US-09-960-706-489	Sequence 489, App
	26	118.2	5.2	472	14	US-10-240-955-193	Sequence 193, App
	27	99.2	4.4	480	12	US-10-198-846-235	Sequence 235, App
	28	99.2	4.4	481	12	US-09-814-333-12715	Sequence 12715, App
	29	96.6	4.3	478	11	US-09-814-333-1751	Sequence 1751, App
	30	96.4	4.3	582	10	US-09-918-995-6431	Sequence 6431, App
	31	96.4	4.3	582	14	US-09-796-632-8536	Sequence 8536, App
	32	88.4	3.9	1848	11	US-10-040-862-8536	Sequence 8536, App
C	33	87.2	3.8	852	14	US-09-945-537-29	Sequence 29, Appl
	34	77	3.4	527	14	US-10-198-846-255	Sequence 255, App
	35	75.4	3.3	785	14	US-10-066-543-2156	Sequence 2156, App
	36	75.4	3.3	2117	10	US-10-102-806-325	Sequence 325, App
	37	75.4	3.3	2117	10	US-09-965-529-53	Sequence 53, Appl
	38	67.4	3.0	2317	14	US-09-969-680A-53	Sequence 53, Appl
	39	62.2	2.7	574	9	US-10-097-340-94	Sequence 94, Appl
	40	62.2	2.7	669	9	US-09-864-761-228	Sequence 228, App
C	41	60.2	2.6	1168	12	US-09-864-761-17051	Sequence 17051, App
	42	60	2.6	60	12	US-10-017-161-2179	Sequence 2179, App
C	43	59.8	2.6	285	9	US-09-908-975-15856	Sequence 15856, App
	44	59.8	2.6	285	11	US-09-737-178-64	Sequence 64, Appl
C	45	59.8	2.6	285	11	US-09-286-488-64	Sequence 64, Appl
						US-09-853-079-64	Sequence 64, Appl

27	99.2	4.4	480	12	US-09-814-333-12715	Sequence 12715, A
28	99.2	4.4	480	12	US-09-814-333-1751	Sequence 1751, Ap
29	99.2	4.4	481	12	US-09-814-333-1751	Sequence 1751, Ap
30	96.6	4.3	478	11	US-09-911-935-6431	Sequence 6431, Ap
31	96.6	4.3	582	10	US-09-796-632-8536	Sequence 8536, Ap
32	96.4	4.3	582	14	US-10-040-862-8536	Sequence 8536, Ap
33	88.4	3.9	1848	11	US-09-945-537-29	Sequence 29, Appl
34	87.2	3.8	852	14	US-10-198-846-255	Sequence 255, App
35	77	3.4	527	14	US-10-065-543-2156	Sequence 2156, Ap
36	75.4	3.3	785	14	US-10-103-806-325	Sequence 325, App
37	75.4	3.3	2117	11	US-09-965-529-53	Sequence 53, Appl
38	62.4	3.0	2317	14	US-10-097-340-94	Sequence 94, Appl
39	62.2	2.7	574	9	US-09-864-761-328	Sequence 288, App
40	62.2	2.7	669	9	US-09-864-761-17051	Sequence 17051, A
41	60.2	2.7	1168	12	US-10-017-161-2179	Sequence 2179, Ap
42	60	2.6	60	12	US-09-908-975-15856	Sequence 15856, A
43	59.8	2.6	285	9	US-09-737-179-64	Sequence 64, Appl
44	59.8	2.6	285	10	US-09-286-488-64	Sequence 64, Appl
45	59.8	2.6	285	11	US-09-853-079-64	Sequence 64, Appl

31	96.4	4.3	582	14	US-10-040-862-8536	Sequence 8536, Ap
32	98.4	4.3	1848	11	US-09-945-537-29	Sequence 29, Appl
33	87.2	3.8	852	14	US-10-198-846-255	Sequence 255, Ap
34	77	3.4	527	14	US-10-066-543-2156	Sequence 2156, Ap
35	75.4	3.3	785	14	US-10-102-806-325	Sequence 325, Ap
36	75.4	3.3	2117	10	US-09-965-529-53	Sequence 53, Appl
37	75.4	3.3	2117	11	US-09-969-680A-53	Sequence 53, Appl
38	67.4	4.0	2317	14	US-10-097-340-94	Sequence 94, Appl
39	62.2	2.7	574	9	US-09-864-761-228	Sequence 228, App
40	62.2	2.7	669	9	US-09-864-761-17051	Sequence 17051, A
41	60.2	2.7	1168	12	US-10-017-161-2179	Sequence 2179, Ap
42	60	2.6	60	12	US-09-908-975-15856	Sequence 15856, A
43	59.8	2.6	285	9	US-09-737-178-64	Sequence 64, Appl
44	59.8	2.6	285	10	US-09-286-488-64	Sequence 64, Appl
45	59.8	2.6	285	11	US-09-853-079-64	Sequence 64, Appl

34	77	527	14	US-10-060-543-2150	Sequence 1586, App
35	75.4	3.3	785	14	Sequence 325, App
36	75.4	3.3	2117	10	Sequence 53, App
37	75.4	3.3	2117	11	Sequence 53, App
38	67.4	3.0	2317	14	Sequence 94, App
39	62.2	2.7	574	9	Sequence 228, App
40	62.2	2.7	669	9	Sequence 17051, A
41	60.2	2.7	1168	12	Sequence 2179, App
42	60	2.6	60	12	Sequence 15856, A
43	59.8	2.6	285	9	Sequence 64, App
44	59.8	2.6	285	10	Sequence 64, App
45	59.8	2.6	285	11	Sequence 64, App

[illegible]

38	67.4	3.0	2317	14	US-10-091-340-54	Sequence 94, Appl
39	62.2	2.7	574	9	US-09-864-761-228	Sequence 228, Appl
40	62.2	2.7	669	9	US-09-864-761-17051	Sequence 17051, A
41	60.2	2.7	1168	12	US-10-017-161-2119	Sequence 2119, A
42	60	2.6	60	12	US-09-908-975-15856	Sequence 15856, A
43	59.8	2.6	285	9	US-09-737-178-64	Sequence 64, Appl
44	59.8	2.6	285	10	US-09-286-488-64	Sequence 64, Appl
45	59.8	2.6	285	11	US-09-853-079-64	Sequence 64, Appl

C 1	60.2	2.7	1168	12	US-10-017-161-2119	Sequence 211/9, App
C 2	60	2.6	60	12	US-09-908-975-15856	Sequence 115/6, A
C 3	59.8	2.6	285	9	US-09-737-178-64	Sequence 64, Appl
C 4	59.8	2.6	285	11	US-09-286-488-64	Sequence 64, Appl
C 5	59.8	2.6	285	11	US-09-853-079-64	Sequence 64, Appl

```

c 45 59.8 2.6 285 11 US-09-853-079-64
Sequence 64, Appl.

ALIGNMENTS

RESULT 1
US-10-171-581-170
; Sequence 170, Application US/10171581
; Publication No. US20030104426A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: Linsley, Peter
; APPLICANT: Mao Mao
; TITLE OF INVENTION: Signature Genes in Chronic Myelogenous Leukemia
; FILE REFERENCE: 9301-157-999
; CURRENT APPLICATION NUMBER: US/10/171,581
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/298,914
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 366
; SEQ ID NO 170
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLIC INFORMATION:
; DATABASE ACCESSION NUMBER: U41060
; DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-170

```

```

? PRIOR FILING DATE: 2001-06-18
? NUMBER OF SEQ ID NOS: 366
? SEQ ID NO 170
? LENGTH: 2744
? TYPE: DNA
? ORGANISM: Homo sapiens
? PUBLICATION INFORMATION:
? DATABASE ACCESSION NUMBER: U41060
? DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-170

```

```

? TYPE: DNA
? ORGANISM: Homo sapiens
? PUBLICATION INFORMATION:
? DATABASE ACCESSION NUMBER: U41060
? DATABASE ENTRY DATE: 2001-06-18
US-10-171-581-170

Query Match          98.0%; Score 2222; DB 14; Length 2744;
Best local Similarity 99.2%; Pred. No. 0;
Matches 2250; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY      1  ATGGCGAGGAAGTATCTGTAATCTTGATCTGACCTTGGCCCTCTCTGTCAAAATCCC 60
        |||||
Db      138  ATGGCGAGGAAGTATCTGTAATCTTGATCTGACCTTGGCCCTCTCTGTCAAAATCCC 197

QY      61  CTTCTGTAAGTAAGACGAGCTGCTTTCCCGACGACCTAGAGAAATTAGTCCGAAATGG 120
        |||||
Db      198  CTTCTGTAAGTAAGACGAGCTGCTTTCCCGACGACCTAGAGAAATTAGTCCGAAATGG 257

QY      121  GAATCTGGCATTAAATGTTGACTTTGGCAATTTCCACAGGCAATATCATCTCAACAGCTT 180
        |||||
Db      258  GAATCTGGCATTAAATGTTGACTTTGGCAATTTCCACAGGCAATATCATCTCAACAGCTT 317

```



```

; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysecul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 263
; LENGTH: 2744
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-177-293-263

Query Match      98.0%; Score 2222; DB 14; Length 2744;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2250; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY      1  ATGGCGAGAGTATCTGTAATCTTGATCTGACCTTTGCCCTCTCTGTCAACAAATCCC 60
DB      138  ATGGCGAGAGTATCTGTAATCTTGATCTGACCTTTGCCCTCTCTGTCAACAAATCCC 197

QY      61  CTTCAAGCACTAAAGCAGCTGCTTTCCCGCAGACCACTAGAAAAATTAGTCCGAATGG 120
DB      198  CTTCAAGCACTAAAGCAGCTGCTTTCCCGCAGACCACTAGAAAAATTAGTCCGAATGG 257

QY      121  GAATCTGACATTAATGTGACTTGGCAATTTCCACAGCGCAATATCATCTCAACAGGTT 180
DB      258  GAATCTGCAATTAATGTGACTTGGCAATTTCCACAGCGCAATATCATCTCAACAGGTT 317

QY      181  TTCACACCGTATGAGAAAAATAATTCCTTTGTGTCAGTTGAAGGGTTCAGAAAAATTACTCAA 240
DB      318  TTCACACCGTATGAGAAAAATAATTCCTTTGTGTCAGTTGAAGGGTTCAGAAAAATTACTCAA 377

QY      241  AATATAGCATAGATAGATTAAGAAATCCATATACCATGACCACTAGCAACCACTCACTCA 300
DB      378  AATATAGCATAGATAGATTAAGAAATCCATATACCATGACCACTAGCAACCACTCACTCA 437

QY      301  GACCACGAGCACTACTCAGACCACTAGAGGCTCAGCTCAGACCACTAGAGGCTCAGACCA 360
DB      438  GACCACGAGCACTACTCAGACCACTAGAGGCTCAGCTCAGACCACTAGAGGCTCAGACCA 497

QY      361  GAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTGCTTCTGGTAAAAAT 420
```

```

Db      498  GAGCATCACTCTG-----ACCATATCATGCTGCTTCTGGTAAAAAT 539
QY      421  AAGCGAAAAAGCTCTTTGCCAGAACCATGACTCAGATAGTTCCAGTAAAGATCCCTAGAAAC 480
Db      540  AAGCGAAAAAGCTCTTTGCCAGAACCATGACTCAGATAGTTCCAGTAAAGATCCCTAGAAAC 599
QY      481  AGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGGTAGAAGGAATGTCAAGGAC 540
Db      600  AGCCAGGGGAAAGAGCTCACCGACCAAGACATGCCAGTGGTAGAAGGAATGTCAAGGAC 659
QY      541  AGTGTAGTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 600
Db      660  AGTGTAGTCTAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 719
QY      601  TTTCTAGAGACAATAGAGACTCCAGAGCTCGGAAAGTCTTCCCAAGATGTAAAGCAGC 660
Db      720  TTTCTAGAGACAATAGAGACTCCAGAGCTCGGAAAGTCTTCCCAAGATGTAAAGCAGC 779
QY      661  TCCACTCCACCCAGTGTCAATCAAGAGCGGGTGGAGCGGCTGGCTGGTAGGAAACA 720
Db      780  TCCACTCCACCCAGTGTCAATCAAGAGCGGGTGGAGCGGCTGGCTGGTAGGAAACA 839
QY      721  AATGAATCTGTGAGTGAGCCCGAAAGGCTTATGTTCCAGAAACACAAATGAAAT 780
Db      840  AATGAATCTGTGAGTGAGCCCGAAAGGCTTATGTTCCAGAAACACAAATGAAAT 899
QY      781  CTTAGAGAGTGTTCATCAATGATCAAGCTACTGATCTCATGCGGATGGGATGGGATGG 840
Db      900  CTTAGAGAGTGTTCATCAATGATCAAGCTACTGATCTCATGCGGATGGGATGGGATGG 959
QY      841  CGGTGATGCAACAGAGTTCACATCTCTGTCGACCACTCATCAACCAATGATGCT 900
Db      960  CGGTGATGCAACAGAGTTCACATCTCTGTCGACCACTCATCAACCAATGATGCT 1019
QY      901  AGATCTTGTCTGATTCATCAAGTGAAGAGGCTGAAATCCCTCCAAAGACCTATTCA 960
Db      1020  AGATCTTGTCTGATTCATCAAGTGAAGAGGCTGAAATCCCTCCAAAGACCTATTCA 1079
QY      961  TTACAATPAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
Db      1080  TTACAATPAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1139
QY      1021  CTGGGGTGTATCTAGTGGCTCTCATGAATCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
Db      1140  CTGGGGTGTATCTAGTGGCTCTCATGAATCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1199
QY      1081  CTTGGCACTGGCCGCTTGGGACTTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1140
Db      1200  CTTGGCACTGGCCGCTTGGGACTTTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1259
QY      1141  TCTCATGCAAGTCAACCACTAGTCAATGACCAAGACCAAGACCAAGACCAAGACCAAGAC 1200
Db      1260  TCTCATGCAAGTCAACCACTAGTCAATGACCAAGACCAAGACCAAGACCAAGACCAAGAC 1319
QY      1201  GGACCACTTTTTCAGTCTGCTCTTCAAAACATAGAAGAAAGTGGCTATTTTGAATTC 1260
Db      1320  GGACCACTTTTTCAGTCTGCTCTTCAAAACATAGAAGAAAGTGGCTATTTTGAATTC 1379
QY      1261  AGTGGAAAGGTCTACAGCTCTAGAGGCTGCTGATTTTCATGTTGTTGTTGTTGTTGTTGTTG 1320
Db      1380  AGTGGAAAGGTCTACAGCTCTAGAGGCTGCTGATTTTCATGTTGTTGTTGTTGTTGTTGTTG 1439
QY      1321  CTCACATTTGATCAACAAATTAAGATAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1380
Db      1440  CTCACATTTGATCAACAAATTAAGATAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1499
QY      1381  GATGATGATGGAGATTAAGAGAGGTTGTCCAAAGTATGATCTCAACTTTCACAAAT 1440
Db      1500  GATGATGATGGAGATTAAGAGAGGTTGTCCAAAGTATGATCTCAACTTTCACAAAT 1559
QY      1441  GAGGAGAAAGTATGATCAAGATGATCGAACTGAAGGCTATTTACGAGAGACTCAAGAG 1500
```

```
Db 1560 GAGGAGAAAGTAGATACAGATGATCGAAGTGAAGGCTATTTACAGAGAGACTCACAAAGAG 1619
QY 1501 CCTCCCACTTTGATTTCTCAGACGCTGCAGTCTTGGAGAGAGAGAGTCAATGATAGCT 1560
Db 1620 CCTCCCACTTTGATTTCTCAGACGCTGCAGTCTTGGAGAGAGAGAGTCAATGATAGCT 1679
QY 1561 CAGTCTCATCCACAGAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1680 CATGCTCATCCACAGAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1739
QY 1621 CATTCACATTTCCAGATACATCGCCAGTGCAGAGTGCAGAGTGCATTCACACCATCATGAC 1680
Db 1740 CATTCACATTTCCAGATACATCGCCAGTGCAGAGTGCATTCACACCATCATGAC 1799
QY 1681 TACATCATATTTCTCCATCATCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1740
Db 1800 TACATCATATTTCTCCATCATCACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 1859
QY 1741 CGTACTCTCGGAGAGAGTGAAGATGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1800
Db 1860 CGTACTCTCGGAGAGAGTGAAGATGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1919
QY 1801 ATGGGTGATGGCTGCACAAATTCAGCGATGGCTGAGCAATGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1920 ATGGGTGATGGCTGCACAAATTCAGCGATGGCTGAGCAATGCTGCTGCTGCTGCTGCTGCTGCT 1979
QY 1861 GGTATCATCAAGTGGTTAAAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
Db 1980 GGTATCATCAAGTGGTTAAAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2039
QY 1921 TTAGTGTACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 2040 TTAGTGTACTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2099
QY 1981 AATGCAATTCAGACGATGCTGGGCTATCTTGAATGGCAACAGAGATTTTCAATGGTGCAT 2040
Db 2100 AATGCAATTCAGACGATGCTGGGCTATCTTGAATGGCAACAGAGATTTTCAATGGTGCAT 2159
QY 2041 TATGCTGAAGATGTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2160 TATGCTGAAGATGTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2219
QY 2101 GCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2220 GCTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2279
QY 2161 CGCTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTTGGGTTTGGGTTTGGGTTTGGGTT 2220
Db 2280 CGCTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTTGGGTTTGGGTTTGGGTTTGGGTT 2339
QY 2221 CTATTTTCCATATTTGACATATAAATTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2268
Db 2340 CTATTTTCCATATTTGACATATAAATTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2387
```

```
RESULT 3
US-10-198-846-10007
; Sequence 10007, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198, 846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10007
; LENGTH: 3537
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3533, 3534, 3535, 3536, 3537
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10007
```

```
Query Match 97.98; Score 2220.4; DB 14; Length 3537;
Best Local Similarity 99.28; Pred. No. 0;
Matches 2249; Conservative 0; Mismatches 1; Indels 18; Gaps 1;

QY 1 ATGGCGAGGAGTATCTGTAATCTTGATCTTCCTGACCTTGCCCTCTCTCTCTCTCTCTCTCTCT 60
Db 138 ATGGCGAGGAGTATCTGTAATCTTGATCTTCCTGACCTTTGCCCTCTCTCTCTCTCTCTCTCT 197
QY 61 CTTTCATGAACTAAAGAGAGTCTTTCCCGCAGACCACTGAGAAAATTAAGTCCGAATGG 120
Db 198 CTTTCATGAACTAAAGAGAGTCTTTCCCGCAGACCACTGAGAAAATTAAGTCCGAATGG 257
QY 121 GAATCTGGCAATTAATGTTGCTTGGCAATTTCCACACGGCAATATCATCTPACAAAGGTT 180
Db 258 GAATCTGGCAATTAATGTTGCTTGGCAATTTCCACACGGCAATATCATCTPACAAAGGTT 317
QY 181 TTTTCACCGTATGAGAGAAAATTAATCTTTGTCAGTTGAAGGTTTCAGAAAATTAAGTCTCAA 240
Db 318 TTTTCACCGTATGAGAGAAAATTAATCTTTGTCAGTTGAAGGTTTCAGAAAATTAAGTCTCAA 377
QY 241 AATATAGGCTAGATATAAGATTAAGAAATCCATATACACCAATGACCAACGACCACTCACTCA 300
Db 378 AATATAGGCTAGATATAAGATTAAGAAATCCATATACACCAATGACCAACGACCACTCACTCA 437
QY 301 GACCAAGGACATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATGAGCATGAGCATGAG 360
Db 438 GACCAAGGACATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATGAGCATGAGCATGAG 497
QY 361 GACCAATCACTCTGACCATGATCATCTCTCACCATAATCATCTGCTGCTGCTGCTGCTGCTGCT 420
Db 498 GACCAATCACTCTG-----ACCAATATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 539
QY 421 AAGCGAAAAGCTCTTTGCCAGACCATGACTCAGATAGTTTCAGTAAAGATCTCTAGAAAC 480
Db 540 AAGCGAAAAGCTCTTTGCCAGACCATGACTCAGATAGTTTCAGTAAAGATCTCTAGAAAC 599
QY 481 AGCCAGGGGAAGAGAGTCTCAGCCAGCAACATGCCAGTGTAGAGGATGTCAGAGGC 540
Db 600 AGCCAGGGGAAGAGAGTCTCAGCCAGCAACATGCCAGTGTAGAGGATGTCAGAGGC 659
QY 541 AGTGTGTAGTCTAGTGAAGTGAAGTCACTCAACTGTGTACAACTGCTCTCTCTCTCTCTCTCTCT 600
Db 660 AGTGTGTAGTCTAGTGAAGTGAAGTCACTCAACTGTGTACAACTGCTCTCTCTCTCTCTCTCTCT 719
QY 601 TTTTCATAGACAAATAGAGAGTCTCAAGACCTGGAAAATCTTCCCGCAAGATGTAGACGAGC 660
Db 720 TTTTCATAGACAAATAGAGAGTCTCAAGACCTGGAAAATCTTCCCGCAAGATGTAGACGAGC 779
QY 661 TCCACTCCACCCAGTGTACATCAAAAGAGCGGGTGAAGCGGCTGGCTGGCTGGCTGGCTGGCTGG 720
Db 780 TCCACTCCACCCAGTGTACATCAAAAGAGCGGGTGAAGCGGCTGGCTGGCTGGCTGGCTGGCTGG 839
QY 721 AATGAATCTGTAGTGAAGCGGCGGAAAAGGCTTTATGTATTTCCAGAAACCAATGAAAT 780
Db 840 AATGAATCTGTAGTGAAGCGGCGGAAAAGGCTTTATGTATTTCCAGAAACCAATGAAAT 899
QY 781 CTTCAAGGAGTGTTCATGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 840
Db 900 CTTCAAGGAGTGTTCATGATGATCAATGATGATGATGATGATGATGATGATGATGATGATGATG 959
QY 841 CGGCTCAATGCAACAGAGTTCAACTATCTCTCTCAGCGCATCATCAACCAATGATGCT 900
```


Db 960 CCGCTGAATGCAACAGAGCTTCAACTATCTCTGTCCAGCCATCATCAACCAAAATGTATGCT 1019
QY 901 AGATCTTGTCGATTCATACAAAGTGAAGAAGGCTGAATCCCTCCAAAGACCTATTCA 960
Db 1020 AGATCTTGTCGATTCATACAAAGTGAAGAAGGCTGAATCCCTCCAAAGACCTATTCA 1079
QY 961 TTACAAATAGCCCTGGGTGGTGGTTTATAGCCATTTGCATCATCAGTTTCCGTGCTG 1020
Db 1080 TTACAAATAGCCCTGGGTGGTGGTTTATAGCCATTTGCATCATCAGTTTCCGTGCTG 1139
QY 1021 CTGGGGTATCTTAGTGCCCTCATGAATCGGGTGTCTTCAAAATTCCTCGAGTTTC 1080
Db 1140 CTGGGGTATCTTAGTGCCCTCATGAATCGGGTGTCTTCAAAATTCCTCGAGTTTC 1199
QY 1081 CTGTGGCACTGECCTGTGGGACTTTGAGTGGTGTGATGCTTTTACACCTTCTTCCACAT 1140
Db 1200 CTGTGGCACTGECCTGTGGGACTTTGAGTGGTGTGATGCTTTTACACCTTCTTCCACAT 1259
QY 1141 TCTCATGCAAGTCACACCATAGTCATGACCATGAGCAACAGCAATGGAATGAAGA 1200
Db 1260 TCTCATGCAAGTCACACCATAGTCATGACCATGAGCAACAGCAATGGAATGAAGA 1319
QY 1201 GGACCACTTTTCAGTCATCTCTCTCAAAACATAGAGAAAGTGCTATTTTGATTC 1260
Db 1320 GGACCACTTTTCAGTCATCTCTCTCAAAACATAGAGAAAGTGCTATTTTGATTC 1379
QY 1261 ACSTGGAAGGCTTAACAGCTCTAGAGGCTGTATTTTCATGTTTCTTGTGACATGTC 1320
Db 1380 ACSTGGAAGGCTTAACAGCTCTAGAGGCTGTATTTTCATGTTTCTTGTGACATGTC 1439
QY 1321 CTCACATGATCAACAATTTAAAGATAAGAGAAAGAAATCAGAGAACTCGAAAT 1380
Db 1440 CTCACATGATCAACAATTTAAAGATAAGAGAAAGAAATCAGAGAACTCGAAAT 1499
QY 1381 GATGATGATGAGATGAAGACGATTTGCCAAGTATGATATCTCACTTTCAACAAAT 1440
Db 1500 GATGATGATGAGATGAAGACGATTTGCCAAGTATGATATCTCACTTTCAACAAAT 1559
QY 1441 GAGGAGAAAGTATACAGATGATCGACATGAGGCTATTTACGAGCAGACTCACAGAG 1500
Db 1560 GAGGAGAAAGTATACAGATGATCGACATGAGGCTATTTACGAGCAGACTCACAGAG 1619
QY 1501 CCTCCACCTTTGATTCACAGCAGCTGCTGCTTGGAGAGAAAGAGGTCAATGATCT 1560
Db 1620 CCTCCACCTTTGATTCACAGCAGCTGCTGCTTGGAGAGAAAGAGGTCAATGATCT 1679
QY 1561 CATGCTATCCACAGAAAGTCTACAATGAATATGTTACCAAGGGTGCAAGATAAATGC 1620
Db 1680 CATGCTATCCACAGAAAGTCTACAATGAATATGTTACCAAGGGTGCAAGATAAATGC 1739
QY 1621 CATTCACATTTCCAGATACACTCGGCAGTCAGAGGATCTCATTCACCACTCATGAC 1680
Db 1740 CATTCACATTTCCAGATACACTCGGCAGTCAGAGGATCTCATTCACCACTCATGAC 1799
QY 1681 TACCATCATTTCCATCATCACACCCACCAAAACCCATTCCTCAGTFCACAGCCAG 1740
Db 1800 TACCATCATTTCCATCATCACACCCACCAAAACCCATTCCTCAGTFCACAGCCAG 1859
QY 1741 CGCTACTCTCGGAGAGGCTGAAGAGATGCCGGGCTGCCACTTTGGCCCTGGATGGTATA 1800
Db 1860 CGCTACTCTCGGAGAGGCTGAAGAGATGCCGGGCTGCCACTTTGGCCCTGGATGGTATA 1919
QY 1801 ATGGGTATGGCCTGCACAAATTCAGGATGGCTAGCAATTTGGTCTGCTTTTACTGAA 1860
Db 1920 ATGGGTATGGCCTGCACAAATTCAGGATGGCTAGCAATTTGGTCTGCTTTTACTGAA 1979
QY 1861 GGCCTATCAAGTGGTAAAGTACTTCTTGTCTGTCTGTCTGATGAGTGGCTCATGAA 1920
Db 1980 GGCCTATCAAGTGGTAAAGTACTTCTTGTCTGTCTGTCTGATGAGTGGCTCATGAA 2039
QY 1921 TTAGGTGACCTTTCGCTGCTTACTAAAGGCTGGCATGACCGTTAAAGGCTGCTCTTAT 1980
Db 2040 TTAGGTGACCTTTCGCTGCTTACTAAAGGCTGGCATGACCGTTAAAGGCTGCTCTTAT 2099

QY 1981 AATGCATTTCCAGCATGCTGGGCTATCTTGGAAATGGCAACAGAAATTTTCATTTGTCAT 2040
Db 2100 AATGCATTTCCAGCATGCTGGGCTATCTTGGAAATGGCAACAGAAATTTTCATTTGTCAT 2159
QY 2041 TATGCTGAAATGTTTCTATGTGGATATTTGCACCTTACTGCTGCTTATTCATGATGTT 2100
Db 2160 TATGCTGAAATGTTTCTATGTGGATATTTGCACCTTACTGCTGCTTATTCATGATGTT 2219
QY 2101 GCTCTGTTTGAATATGTTACCTGAAATGCTGCACAAATGATGCTAGTACCATGGATGATG 2160
Db 2220 GCTCTGTTTGAATATGTTACCTGAAATGCTGCACAAATGATGCTAGTACCATGGATGATG 2279
QY 2161 CGTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATGTTA 2220
Db 2280 CGTGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATGTTA 2339
QY 2221 CTTATTTCCATATTTGAACTAAATCGTGTTCGTATATAAATTTCTAG 2268
Db 2340 CTTATTTCCATATTTGAACTAAATCGTGTTCGTATATAAATTTCTAG 2387

RESULT 4
US-10-176-847-51
; Sequence 51, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Velby, Pether Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREA
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREA
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 3586
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-176-847-51

Query Match 96.9%; Score 2198; DB 14; Length 3586;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2248; Conservative 0; Mismatches 0; Indels 20; Gaps 3;

QY 1 ATGCGAGGAAGTTATCTGTAATCTTGATCCTGACCTTTGGCCCTCTCTGTCAACAATCC 60
Db 263 ATGCGAGGAAGTTATCTGTAATCTTGATCCTGACCTTTGGCCCTCTCTGTCAACAATCC 322
QY 61 CTTTCATGAACCTAAAGCAGCTGCTTTCCCGACAGCAGCTGAGAAATTTAGTCCGAATTGG 120
Db 323 CTTTCATGAACCTAAAGCAGCTGCTTTCCCGACAGCAGCTGAGAAATTTAGTCCGAATTGG 382
QY 121 GAATCTGGCAATTAATGTTGACTTGGCAATTTCCACAGGCAATATCATCTACACAGCTT 180
Db 383 GAATCTGGCAATTAATGTTGACTTGGCAATTTCCACAGGCAATATCATCTACACAGCTT 442
QY 181 TTCTACCGGTATGAGAGAAATTAATCTTTGTCTAGTTGAGGGTTTCAGAAAATTTACTTCA 240
Db 443 TTCTACCGGTATGAGAGAAATTAATCTTTGTCTAGTTGAGGGTTTCAGAAAATTTACTTCA 502
QY 241 AATATAGGCATAGATAAGATTAAGAAATCCATATACACCATGACACCATCATCTCA 300
Db 503 AATATAGGCATAGATAAGATTAAGAAATCCATATACACCATGACACCATCATCTCA 562
QY 301 GACCACGAGCATCACTCAGAACCATGAGCGTCACTCAGAACCATGAGCGATCACTCAGACCC 360
Db 563 GACCACGAGCATCACTCAGAACCATGAGCGTCACTCAGAACCATGAGCGATCACTCAGACCC 622
QY 361 GAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTGCTTCTGGTAAATAT 420
Db 623 GAGCATCACTCTGACCATGATCATCACTCTCACCATAATCATGCTGCTTCTGGTAAATAT 664

; ORGANISM: Homo sapiens
US-09-925-300-591

```
Query Match      50.3%; Score 1141.4; DB 10; Length 1193;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1139; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 539 ACAGTGTATAGTCTAGTGAAGTACCTCAACTGTGTACAACTGTCTCTGAAGGAATC 598
Db 1 ACAGTGTATAGTCTAGTGAAGTACCTCAACTGTGTACAACTGTCTCTGAAGGAATC 60

QY 599 ACTTCTAGACACATAGACACTCCAGACCTGGAAACTCTTCCCAAGATGTAGCA 658
Db 61 ACTTCTAGACACATAGACACTCCAGACCTGGAAACTCTTCCCAAGATGTAGCA 120

QY 659 GCTCCACTCCACCAGTGTACATCAAAAGACCGGGTGGCGGTGGTGTAGGAAA 718
Db 121 GCTCCACTCCACCAGTGTACATCAAAAGACCGGGTGGCGGTGGTGTAGGAAA 180

QY 719 CAAATGAATCTGTAGTGAGCCCGGAAAGGCTTTATGTATTCAGAGAAACAAATGAAA 778
Db 181 CAAATGAATCTGTAGTGAGCCCGGAAAGGCTTTATGTATTCAGAGAAACAAATGAAA 240

QY 779 ATCCTCAGAGGTGTTCAATGCATCAAGCTACTGACATCTCATGGCATGGCATCCAG 838
Db 241 ATCCTCAGAGGTGTTCAATGCATCAAGCTACTGACATCTCATGGCATGGCATCCAG 300

QY 839 TTCGGTGAATGCAACAGAGTCAACTPATCTGTCCAGCCATCATCAACCAAAATGATG 898
Db 301 TTCGGTGAATGCAACAGAGTCAACTPATCTGTCCAGCCATCATCAACCAAAATGATG 360

QY 899 CTAGACTTGTCTGATTCATACAGTGAAGAGGCTGAATCCCTCCAAAGACCTATT 958
Db 361 CTAGACTTGTCTGATTCATACAGTGAAGAGGCTGAATCCCTCCAAAGACCTATT 420

QY 959 CATTCAAAATAGCTGGGTGTTTATAGCCATTTCATCATCAGTTTCTCTCTC 1018
Db 421 CATTCAAAATAGCTGGGTGTTTATAGCCATTTCATCATCAGTTTCTCTCTC 480

QY 1019 TGCTGGGGTGTATCTTATGCTCTCATGAATCGGGTGTATTCAAATTTCTCCTGAGTT 1078
Db 481 TGCTGGGGTGTATCTTATGCTCTCATGAATCGGGTGTATTCAAATTTCTCCTGART 540

QY 1079 TCCTTGTGCACTGCCGTGTGGACTTTGAGTGGTGATGCTTTTACACTTCTCCAC 1138
Db 541 YCCTTGTGCACTGCCGTGTGGACTTTGAGTGGTGATGCTTTTACACTTCTCCAC 600

QY 1139 ATTCTCATCAAGTCAACCAATAGTCATAGCCATGAAGAACCAACCAATGAAATGAAA 1198
Db 601 ATTCTCATCAAGTCAACCAATAGTCATAGCCATGAAGAACCAACCAATGAAATGAAA 660

QY 1199 GAGGACACTTTTCAGTCACTGTCTCTCAAAACATAGAGAAAGTGCCTATTTTGATT 1258
Db 661 GAGGACACTTTTCAGTCACTGTCTCTCAAAACATAGAGAAAGTGCCTATTTTGATT 720

QY 1259 CCAGTGGAGGGTCTAAAGCTCTAGAGGCGCTGATTTCATGTTTCTGTTCAACATG 1318
Db 721 CCAGTGGAGGGTCTAAAGCTCTAGAGGCGCTGATTTCATGTTTCTGTTGACATG 780

QY 1319 TCCTCAGATGTATCAACAAATTTAAAGATGAAGAAAGAAAGATCAGAGAAACCTGAAA 1378
Db 781 TCCTCAGATGTATCAACAAATTTAAAGATGAAGAAAGAAAGATCAGAGAAACCTGAAA 840

QY 1379 ATGATGATGATGTGAGATTAAGAGCAGTTGTCACAGTATGATCTCACTTTCACAA 1438
Db 841 ATGATGATGATGTGAGATTAAGAGCAGTTGTCACAGTATGATCTCACTTTCACAA 900

QY 1439 ATGAGGAGAAAGTATGATACATGATCGAACTGAAGGCTATTATGAGCAGAGTCAAG 1498
Db 901 ATGAGGAGAAAGTATGATACATGATCGAACTGAAGGCTATTATGAGCAGAGTCAAG 960

QY 1499 AGCCCTCCACTTGTATCTCAGCAGGCTGTGAGAGAGAGAGGTGATGATGAT 1558
Db 1558 AGCCCTCCACTTGTATCTCAGCAGGCTGTGAGAGAGAGAGGTGATGATGAT
```

```
Db 961 AGCCCTCCACTTTGATTTCTCAGCAGCCTGCAGTCTTGGAAAGAAAGAGGTCATGATAG 1020
QY 1559 CTCACTCTCATCCACAGGAAGTCTACAATGAATATGTATCCACAGAGGGTGCAGAAATATAT 1618
Db 1021 CTCACTCTCATCCACAGGAAGTCTACAATGAATATGTATCCACAGAGGGTGCAGAAATATAT 1080
QY 1619 GCAATTCATATTCACAGATACACTCGGCCAGTCAGAGATCTCAATTCACACCAATCATG 1678
Db 1081 GCAATTCATATTCACAGATACACTCGGCCAGTCAGAGATCTCAATTCACACCAATCATG 1140
QY 1679 ACT 1681
Db 1141 ACT 1143

RESULT 6
US-10-198-846-10321
; Sequence 10321, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steimann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; FILE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10321
; LENGTH: 1449
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1447, 1448, 1449
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10321

Query Match      39.7%; Score 900.4; DB 14; Length 1449;
Best Local Similarity 96.4%; Pred. No. 8e-248;
Matches 985; Conservative 0; Mismatches 31; Indels 6; Gaps 6;

QY 967 ATAGCCTGGGTGGTGGTGTATAGCCATTCCATCATCAGTTTCTCTGCTGCTGGGG 1026
Db 199 AGAGCCTGGGTGGTGGTGTATAGCCATTCCATCATCAGTTTCTCTGCTGCTGGGG 258

QY 1027 GTATCTTAGTGCCCTCAGTAATGCGGGTGTATTCCTCAAAATTCCTCTGAGTTTCTCTGG 1086
Db 259 GTATCTTAGTGCCCTCAGTAATGCGGGTGTATTCCTCAAAATTCCTCTGAGTTTCTCTGG 318

QY 1087 GCATCGCGGTGGGACTTTGAGTGGTGATGCTTTTACACCTCTCTTCCATCTCAT 1146
Db 319 GCATCGCGGTGGGACTTTGAGTGGTGATGCTTTTACACCTCTCTTCCATCTCAT 378

QY 1147 GCAAGTCACCACTAGTGTATAGCCATGAAGAACCAAGCAATGGAATGAAAGAGGACCA 1206
Db 379 GCAAGTCACCACTAGTGTATAGCCATGAAGAACCAAGCAATGGAATGAAAGAGGACCA 438

QY 1207 CTTTTCAGTCACTCTCTCTCAAAACATAGAGAAAGTGCCTATTTTGTATTCACGTGG 1266
Db 439 CTTTTCAGTCACTCTCTCTCAAAACATAGAGAAAGTGCCTATTTTGTATTCACGTGG 498

QY 1267 AAGGCTCAACAGCTCAGAGGCGGTGATTTTCATGTTTCTTGTGACATGCTCTACA 1326
Db 499 AAGGCTCAACAGCTCAGAGGCGGTGATTTTCATGTTTCTTGTGACATGCTCTACA 558

QY 1327 TTGATCAAAACAATTTAAAGATAAGAAAGAAAGAAAGTCAAGAAACCTGAAATGATGAT 1386
```

Db	559	TTGATCACAACRAATTTAAGAGATAGAGAGAGAAAAGAAATCAGAAGAAACCTGAAATGATGAT	618
QY	1387	GATGTGGAGATTAAAGACGATGTGTCCAAAGTAGTAATCAACTTTCAACAATGAGGAG	1446
Db	619	GATGTGGAGATTAAAGACGATGTGTCCAAAGTAGTAATCAACTTTCAACAATGAGGAG	678
QY	1447	AAAGTAGATACAGATGATCGAAGCTGAAGGCTATTTACGAGCAGACTCAACAGAGCCCTCC	1506
Db	679	AAAGTAGATACAGATGATCGGACACTGAAGGCTATTTACGAGCAGACTCAACAGAGCCCTCC	738
QY	1507	CACTTTGATTTCTCAGCAGCCCTGCAGTGTGGAAGAGAGAGGTCTATGATAGCTCATGCT	1566
Db	739	CACTTTGATTTCTCAGCAGCCCTGCAGTGTGGAAGAGAGAGGTCTATGATAGCTCATGCT	798
QY	1567	CATCCACAGGAGTCTCAANTGATATGTATCCACAGAGGTCAGAGATAAATGCCATTCA	1626
Db	799	CATCCACAGGAGTCTCAATGAATATGTATCCACAGAGGTCAGAGATAAATGCCATTCA	858
QY	1627	CATTTCACAGATACACTCGGCCAGTCAGAGCATCTCATTCACCACCATCATGACTACCAT	1686
Db	859	CATTTCACAGATACACTCGGCCAGTCAGAGCATCTCATTCACCACCATCATGACTACCAT	918
QY	1687	CATATPCTCCATCATCACACCCACCCAAACCCACCATCTCTCAGACTCAGCAGCAGCG-CTA	1745
Db	919	CATATPCTCCATCATCACACCCACCCAAACCCACCATCTCTCAGACTCAGCAGCAGCGACTA	978
QY	1746	CTCTCGSGAGAGCTGAAAGATGCGGCGCTCGCCACTTTGGCCTTGAGTGGTGATATATGGG	1805
Db	979	CTCTCGSGAGAGCTGAAAGATGCGGCGCTCGCCACTCTGGCCCTGGATATATGGG	1038
QY	1806	TGATGSCCTTGCACAA-TTTCAGAGATGGCCTAGCAATTGG--TGCTGCTTTTACTGAA--GG	1862
Db	1039	TGATGSCCTTGCACAAATTTCAGAGATGGCCTAGCAATTGGCTGTGCTGCTTTTACTGAAAGG	1098
QY	1863	CTTATCAAGTTGGTTTAAAGTACTGTGTGTGTCTGTCTCATGAGTTGCCCTCATGAATT	1922
Db	1099	CTTATCAAGTTGGTTTAAAGTACTGTGTGTGTCTGTCTCATGAGTTGCCCTCATGAATT	1157
QY	1923	AGGTGACTTGTGTGTTCTACTTAAAGGTGTCGATGACCCCTTAAGCAGGCTGTGCTTTATAA	1982
Db	1158	AGGAGACTTAGCTGTCTACTTCAAGGGTGGCATTAACG--TAAGCCGGCTGTCTCTTGCA	1216
QY	1983	TG 1984	
Db	1217	TG 1218	

```

RESULT 7
US-10-198-846-13805
; Sequence 13805, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13805
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

```

:	LOCATION:	1004,	1005,	1006,	1007,	1008,	1009,	1010,	1011,	1012,	1013,
:		1014,	1015,	1016,	1017,	1018,	1019,	1020,	1021,	1022,	1023,
:	LOCATION:	1024,	1025,	1026,	1027,	1028,	1029,	1030,	1031,	1032	
:	OTHER INFORMATION:	n = A,T,C or G									
US-10-198-846-13805											
<hr/>											
	Query Match	30.6%; Score 694.2; DB 14; Length 1032;									
	Best Local Similarity	90.0%; Pred. No. 1.3e-188;									
	Matches 800; Conservative	0; Mismatches 88; Indels 6; Gaps 5;									
<hr/>											
QY	214	GTTGAGGGTTCAGAAAAATTACTTCAAAATATATAGCATAGATAAGAATAAAGAAATCCAT	273								
Db	4	GTCGGAAGGGTTAGAAAATTACTT-AAATATATAGCATAGATAAGAATAAAGAAATCCAT	62								
QY	274	ATACACCATGCACCGACCATCTCAGTCCAGCAGAGCATCAC-TCAGACCATTGASC-GTC	331								
Db	63	ATACACCATGCACCGACCATCTCAGACCCAGAGCATCAGCTCAGACCATGAGCGGTCT	122								
QY	332	ACTCAGACCATGAGC-ATCACTCAGACCAGCAGCATCACTTGACCATGATCATCACTCT	390								
Db	123	ACTCAGACCATGAGCAATCACTCAGACCAGCAGCATCACTCTGACCATGATCATCACTCT	182								
QY	391	CACATAATCATGCTGCTTCGTGTAATAATAGCCGAAAGCTCTTTGCCAGACCATGAC	450								
Db	183	CACATAATCATGCTGCTTCGTGTAATAATAGCGAAAGCTCTTTGCCAGACCATGAC	242								
QY	451	TCAGATAGTTTCAGSTAAAGATCCCTAGAAACAGCCAGGGGAAAGGAGCTCACCGAACCGAA	510								
Db	243	TCAGATAGTTTCAGSTAAAGATCCCTAGAAACAGCCAGGGGAAAGGAGCTCACCGACGAA	302								
QY	511	CATGCCAGTGTAGAGGAATGTCAGGCACAGTGTAGTGTCTAGTAGTGAAGTGACCTCAACT	570								
Db	303	CATGCCAGTGTAGAGGAATGTCAGGCAGTGTGTAGTGTAGTGAAGTGACCTCAACT	362								
QY	571	GTGTACAACACTGTCTCTGAGGAACCTCACTTCTTAGAGCAATAGAGACTCCAGAGCT	630								
Db	363	GTGTACAACACTGTCTCTGAGGAACCTCACTTCTTAGAGCAATAGAGACTCCAGAGCT	422								
QY	631	GGAAACTCTTCCCAGAAAGATGTAAGCAGCTCCACTCCACCCAGCTGCACATCAAGAGC	690								
Db	423	GGAAACTCTTCCCAGAAAGATGTAAGCAGCTCCACTCCACCCAGCTGCACATCAAGAGC	482								
QY	691	CGGTGAGCGCGTGGCTGGTAGGAAACAAATGTAATGTGTAGTGAGCCCCGAAAGGC	750								
Db	483	CGGTGAGCGCGTGGCTGGTAGGAAACAAATGTAATGTGTAGTGAGCCCCGAAAGGC	542								
QY	751	TTTATGTAATCCAGAAACACAAATGAAATPCCTCAGGAGTGTTCATGATCAATCAAGCTA	810								
Db	543	TTTATGTAATCCAGAAACACAAATGAAATPCCTCAGGAGTGTTCATGATCAATCAAGCTA	602								
QY	811	GTGACATCTCATGGCATGGGCATCCAGGTTTCGGCTGAATGCAACAGAGTTTCAACTATCTC	870								
Db	603	GTGACATCTCATGGCATGGGCATCCAGGTTTCGGCTGAATGCAACAGAGTTTCAACTATCTC	662								
QY	871	TGTCAGCCATCATCAACCAATGTAGCTGTAGATCTGTCTGATTTCATCAAGTGAAGAG	930								
Db	663	TGTCAGCCATCATCAACCAATGTAGCTGTAGATCTGTCTGATTTCATCAAGTGAAGAG	722								
QY	931	AAGGCTGAATVCTCCAAAGACCTATTCTATFACAAATAGAGCTGGTGGTGGTTTTATA	990								
Db	723	AAGGCTGAATVCTCCAAAGAGCTAATTCATTTCACAAATAGGCTGAGTTGGGTTATATA	782								
QY	991	GCCATFTCCATCATCAGTTTCTCTCTGCTGGGGTTATCTTAGTGCTCTCATGAT	1050								
Db	783	GGCAATTCATCAATCAGTACTG-CACTGCGGGGGTAAATCAAGCGCCCTTAATGAT	840								
QY	1051	CGGGTGGTTTTCAAATTTCTCCTGAGTTTCCCTGTGGGCAGCTGGCGGTG	1099								
Db	841	CGGGGGTTTCCAAATPCCCTGGCATTCCTTGTGACTTGCCTGGCAGCTTG	889								

RESULT 8
US-10-198-846-8490

```
; Sequence 8490, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8490
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1, 44, 516, 571, 600, 606
; OTHER INFORMATION: n = A,T,C or G
```

Query Match 25.5%; Score 578.4; DB 14; Length 606;

Best Local Similarity 99.08; Pred. No. 1.8e-155;

Matches 590; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1042 CTCATGAATCGGGTCTTTTCAAAATTCCTCCAGTCTTCCTGGGCACTGGCCGTGGG 1101
Db 11 CGATGAATCGGGTCTTTTCAAAATTCCTCCAGTCTTCCTGGGCACTGGCCGTGGG 69
QY 1102 ACTTTGAGTGGTATGCTTTTACACCTTCTTCCACATTTCTCATGCAAGTCCACCACT 1161
Db 70 ACTTTGAGTGGTATGCTTTTACACCTTCTTCCACATTTCTCATGCAAGTCCACCACT 129
QY 1162 AGTCATAGCCATGAGAACACAGCAATGGAAATGGAATGAAAGAGGACCACTTTTCACTG 1221
Db 130 AGTCATAGCCATGAGAACACAGCAATGGAAATGGAATGGAAGAGGACCACTTTTCACTG 189
QY 1222 TCTTCCTCAAAACATAGAGAAAGTGCTATTTTGAATTCACGTGGAAGGGTCTTAACAGCT 1281
Db 190 TCTTCCTCAAAACATAGAGAAAGTGCTATTTTGAATTCACGTGGAAGGGTCTTAACAGCT 249
QY 1282 CTAGGAGGCGCTGATTTTCACTGTTCTTGTGTGAACATGTCCTCACTGATFCAACAAATTT 1341
Db 250 CTAGGAGGCGCTGATTTTCACTGTTCTTGTGTGAACATGTCCTCACTGATFCAACAAATTT 309
QY 1342 AAAGATAAGAGAAAGAAATCAGAGAAACCTGAAATGATGATGATGGAGATTAAAG 1401
Db 310 AAAGATAAGAGAAAGAAATCAGAGAAACCTGAAATGATGATGATGGAGATTAAAG 369
QY 1402 AAGCAGTTGTCAGATGATATCTCAACTTTCACAAATGAGGAAAGTAGATACAGAT 1461
Db 370 AAGCAGTTGTCAGATGATATCTCAACTTTCACAAATGAGGAAAGTAGATACAGAT 429
QY 1462 GATCGAGTGAAGGCTATTTACGAGCGAGTCAAGAGCCCTCCACTTTGATTTCTAG 1521
Db 430 GATCGAGTGAAGGCTATTTACGAGCGAGTCAAGAGCCCTCCACTTTGATTTCTAG 489
QY 1522 CAGCCTGCGAGTCTTGGAGAGAGAGAGGTCATGATGATGATGATGATGATGATGATGATGAT 1581
Db 490 CAGCCTGCGAGTCTTGGAGAGAGAGAGGTCATGATGATGATGATGATGATGATGATGATGAT 549
QY 1582 TACAATGAATATGTACCCAGAGGGTGCAAGAAATGCCATTGCCATTGCCATTGCCATTGCCA 1637
Db 550 TACAATGAATATGTACCCAGAGGGTGCAAGAAATGCCATTGCCATTGCCATTGCCA 605
```

RESULT 9

```
US-09-918-995-30283
; Sequence 30283, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; CURRENT APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30283
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-918-995-30283
```

Query Match 14.8%; Score 335.6; DB 11; Length 366;

Best Local Similarity 96.1%; Pred. No. 8.9e-86;

Matches 344; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

```
QY 1671 CCATCATGACTACCATCATATCTCCATCATCCACACACCAAAACCAACCATCTTACAG 1730
Db 9 CCATCATGACTACCATCATATCTCCATCATCCACACACCAAAACCAACCATCTTACAG 68
QY 1731 TCACAGCCAGCGCTACCTCGGAGGAGGTGAAAGATGCCGGCTGCCACATTTGGCTG 1790
Db 69 TCACAGCCAGCGCTACCTCGGAGGAGGTGAAAGATGCCGGCTGCCACATTTGGCTG 128
QY 1791 GATGCTGATATGAGGTGATGGCTGCACAAATTCAGCGATGGCTAGCAATTTGGTGTGC 1850
Db 129 GATGCTGATATGAGGAGGTGCTGCACAAATTCAGCGATGGCTAGCAATTTGGGCTGC 188
QY 1851 TTTTACTGAAGGCTATCAAGTGGTTAAAGTACTTCTGCTGTGTCTGTCTGTCATGAGTT 1910
Db 189 TTTTACTGAAGGCTATCAAGGAGGTAAAGTACTTCTGCTGTGTCTGTCTGTCATGAGTT 248
QY 1911 GCTCATGAATTAGTGACATTTCTCTTCTTACTAAAGCTGGCATGACCGTTAAGCAGGC 1970
Db 249 GCTCATGAATTAGTGACATTTCTGCTGTCTTACTAAAGCTGGCATGACCGTTAAGCAGGC 308
QY 1971 TGTCCTTTTATATGATTCACCATGCTGCGTACTTGGATGCGACAGGATTT 2028
Db 309 TGTCCTTTTATATGATTCACCATGCTGCGTACTTGGATGCGACAGGATTT 366
```

RESULT 10

US-10-198-846-7874

; Sequence 7874, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; PRIOR FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7874

; LENGTH: 905

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

```
; NAME/KEY: misc_feature
; LOCATION: 4, 583, 690, 713, 714, 721, 731, 756, 787, 788, 790, 792.
; LOCATION: 800, 802, 804, 811, 816, 832, 838, 840, 844, 854, 859, 873,
; LOCATION: 875, 887, 888, 898
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-7874

Query Match      14.6%; Score 331.8; DB 14; Length 905;
Best Local Similarity 98.0%; Pred. No. 2.2e-84;
Matches 336; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 246 AGGCATAGTAAGATTAAAGAAATCCATATACACCATGACCCAGCACCATCACTCAGACCA 305
DB 73 AGGCATAGTAAGATTAAAGAAATCCATATACACCATGACCCAGCACCATCACTCAGACCA 132
QY 306 CGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCACTCAGACCATGAGCA 365
DB 133 CGAGCATCACTCAGACCATGAGCGTCACTCAGACCATGAGCATCACTCAGACCATGAGCA 192
QY 366 TCACCTCGACCATGATCATCACTCTCCACCATATCACTGCTCTGCTGTAATATAGCG 425
DB 193 TCACCTCGACCATGATCATCACTCTCCACCATATCACTGCTCTGCTGTAATATAGCG 252
QY 426 AAAAGCTCTTTGCCAGACCATGACTCAGATAGTTCAGGTAAAGATCCCTAGAAACAGCCA 485
DB 253 AAAAGCTCTTTGCCAGACCATGACTCAGATAGTTCAGGTAAAGATCCCTAGAAACAGCCA 312
QY 486 GGGGAAGGAGCTCAGCCAGACATGCGAGTGGGTAGAGGATGTCAAGGACAGTGT 545
DB 313 GGGGAAGGAGCTCAGCCAGACATGCGAGTGGGTAGAGGATGTCAAGGACAGTGT 372
QY 546 TAGTCTAGTGAAGTGAAGTCACTCACTGATGATGATGATGATGATGATGATGATGATGAT 588
DB 373 TAGTCTAGTGAAGTGAAGTCACTCACTGATGATGATGATGATGATGATGATGATGATGAT 415

RESULT 11
US-09-920-455-250
; Sequence 250, Application US/09920455
; Patent No. US20020168647A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF HEAD AND NECK CANCER
; FILE REFERENCE: 210121.540
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 250
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 42, 43
; OTHER INFORMATION: n = A,T,C or G
US-09-920-455-250

Query Match      14.1%; Score 319.2; DB 10; Length 337;
Best Local Similarity 98.5%; Pred. No. 4.4e-81;
Matches 321; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1333 AAACATTTAAAGATAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATG 1392
DB 12 AATTCATTTAAAGATAAGAAAGAAAGATCAGAGAAACCTGAAATGATGATGATG 71
QY 1393 GAGATTAGAGCAGTGTGTCACCATGATGATGATGATGATGATGATGATGATGATG 1452
DB 72 GAGATTAGAGCAGTGTGTCACCATGATGATGATGATGATGATGATGATGATGATG 131
QY 1453 GATACAGATGATGAACTGAAGGCTATTATACGAGCAGACTCACAAGAGCCCTCCACTTT 1512
```

```
DB 132 GATACAGATGATCGAAGCTGAAGGCTATTATACGAGCAGACTCACAAGAGCCCTCCACTTT 191
QY 1513 GATTCTCAGCAGCCTGCGAGTCTTGGAAAGAGAGAGGTGATGATGATGATGATGATGATG 1572
DB 192 GATTCTCAGCAGCCTGCGAGTCTTGGAAAGAGAGAGGTGATGATGATGATGATGATGATG 251
QY 1573 CAGGAAGTCTCAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1632
DB 252 CAGGAAGTCTCAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 311
QY 1633 CACGATACACTCGGCAGCTCAGACGA 1658
DB 312 CACGATACACTCGGCAGCTCAGACGA 337

RESULT 12
US-09-814-353-22020
; Sequence 22020, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22020
; LENGTH: 4417
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4415, 4416, 4417
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-22020

Query Match      14.1%; Score 319; DB 12; Length 4417;
Best Local Similarity 56.3%; Pred. No. 3.4e-80;
Matches 736; Conservative 0; Mismatches 535; Indels 36; Gaps 6;

QY 969 AGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1028
DB 561 AGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 620
QY 1029 TATCTTAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1088
DB 621 GATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 680
QY 1089 ACTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1148
DB 681 ATTAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 740
QY 1149 AAGTCACCCACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1208
DB 741 TGGACATGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATCAGATC 785
```



```
QY 1626 ACATTTCACGATACACTGGCCAGTCAGAGCATCTCATTCACACCATCATGATACCA 1685
Db 2006 -CATGAGCATGATCTCCATGCTGCGCATCAATACCAACCGCGGAGAACAAATCTGCT 2064
QY 1686 TCATATTCTCCATCATCACACCCACCAAAACCACTCTCATCATGATACACGAGGCTA 1745
Db 2065 GAGGAGCATTAATCACCAGTCGCCACACAGCATCTCATCATGATGCGCCCTGTCA 2124
QY 1746 CTCCTCGGAGAGCTGAAGAAGTCGGCGTCGCCCATCTTGGCTCGAATGGATGAATGG 1805
Db 2125 TTCTGGATCCGATCTGAAAGAACAGGAATAGTAATATAGCCTGGATGATCATCGG 2184
QY 1806 FGATGGCTGCACAAATTCAGCGATGGCTAGCAATTTGCTGCTTTTACTGAAGCTT 1865
Db 2185 GGATGGCATCCCAACTTCAGTGAATGGCTCGCAATTTGGTCAGCTTTCAGTCTGGAT 2244
QY 1866 APCAAGTGGTTTAAGTACTTCTTGCTGTGTTCTGTCATGATGAGTTCCTCATGAATPAG 1925
Db 2245 GACAGGAGGAATCAGTACTTCTATAGCGGCTTCTGTCATGAATGCCACATGAATAGG 2304
QY 1926 TGACTTGTCTTACTAAAGCTGGCAGTACGAGCTTAAGCAGGCTGCTCTTATATATGC 1985
Db 2305 AGATTTGGCAGTTCTTCTTAAAGCAGGATGACTGTAAAGCAAGCAATTTGATACACCT 2364
QY 1986 ATTTCTCAGCATGCTGCGCTATCTTGGAAATGGCAACAGGAATTTTCATGCTCATATGC 2045
Db 2365 CCTCTCTGCCATGATGGCTTACATAGGCATGCTCATAGGACACAGCTGTTGGTCAGTATGC 2424
QY 2046 TGAATAGTTTCTATGTGATATTTGCACTTACTGCTGCTTATTCATGATGATGCTCT 2105
Db 2425 CAATAAATCATCACTTTGAGTCTTTCAGTCACTGCGAGCATGTTCTCTATGAGCTT 2484
QY 2106 GATTGATATGTGACTTGAATGCTGCACAAATGAG-----CTAGTGACCATGATGTAG 2159
Db 2485 GGTGATATGCTTCCAGAAATGTTCCATGCTGATGATGATGATGATGATGATGATG 2544
QY 2160 CGCTGGGGGATTTCTTTTACAGAAATGCTGGGATGCTTTGGGTTTGGAAATATGTT 2219
Db 2545 TCGTGGGGCAATTCATCTTCAGAAATTTAGGATTTGCTTTGGATTTGCCATATGCT 2604
QY 2220 ACTTATTCATATTTGAACATAAATCGTGTTCGTATATAATTTCT 2266
Db 2605 GGTGATGGCCCTCTATGAAGATAAATTTGTTGTGACATCCAGTTTT 2651
```

```
RESULT 14
US-10-198-846-9288/c
; Sequence 9288, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9288
; LENGTH: 721
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8, 11, 42, 262, 279, 324, 333, 337, 344, 346, 349, 352, 374,
; LOCATION: 381, 383, 399, 400, 401, 415, 417, 427, 430, 431, 432, 439,
; LOCATION: 445, 454, 457, 458, 459, 464, 468, 469, 470, 473, 475, 477,
```

```
; LOCATION: 488, 493, 495, 504, 509, 519, 529, 536, 545, 547
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 552, 555, 559, 560, 564, 565, 578, 580, 581, 585, 587, 589,
; LOCATION: 590, 593, 594, 600, 604, 609, 611, 623, 625, 626, 627, 628,
; LOCATION: 629, 642, 643, 644, 647, 648, 651, 658, 664, 671, 673, 676,
; LOCATION: 679, 681, 691, 694, 695, 696, 698, 703, 706, 712, 713
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 714, 718
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-9288

Query Match 11.8%; Score 268; DB 14; Length 721;
Best Local Similarity 89.1%; Pred. No. 4.1e-66;
Matches 294; Conservative 0; Mismatches 35; Indels 1; Gaps 1;

QY 1350 TCATGATAGTCATGCTCATCCACGGAAGTCTACAAATATGTATCCACGAGGTCGA 1609
Db 365 TCCTAAAACCCGNGNATNCNTAGTTTANACGGCCNAGGTACCCAGAGGTCGA 306
QY 1610 AGATAAATGSCCATTCACATTTCCACG-AFACACTGGCCAGTCAGACATCTCATAC 1668
Db 305 AGATAAATGCCATTCACATTTCCACNAATACACTGGCCAGTNAAGCATCTCATTCAC 246
QY 1669 CACATCATGATACCATCATATTTCTCCATCATCACACCAACCAACCACTCTCAC 1728
Db 245 CACATCATGATACCATCATATTTCTCCATCATCACACCAACCAACCACTCTCAC 186
QY 1729 AGTCACAGCCAGCGTACTCTCGGAGAGAGTGAAGATGCGGCGCTGGCCACTTTGGCC 1788
Db 185 AGTCACAGCCAGCGTACTCTCGGAGAGAGTGAAGATGCGGCGCTGGCCACTTTGGCC 126
QY 1789 TGGATGTGATATGAGTGGTGTGATGGCTTCACAAATTTCAGCGATGCGCTAGCAATTTGGTCT 1848
Db 125 TGGATGTGATATGAGTGGTGTGATGGCTTCACAAATTTCAGCGATGCGCTAGCAATTTGGTCT 66
QY 1849 GCTTTTACTGAAGGCTTATCAAGTGGTTTA 1878
Db 65 GCTTTTACTGAAGGCTTATCAAGNGTTTAA 36

RESULT 15
US-10-198-846-1455/c
; Sequence 1455, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1455
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2, 3, 5, 7, 9, 69, 211, 224, 360, 405, 435, 454, 488, 493
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1455
```

Query Match	11.28;	Score 253;	DB 14;	Length 504;
Best Local Similarity	87.28;	Pred. No. 6.6e-62;		
Matches 389;	Conservative 0;	Mismatches 42;	Indels 15;	Gaps 10;
QY	1168	AGCCATGAGACACCAAGCAATGGAAATGAAAGAGAGACCACTTTTCAGTCA-TCTGTCTTC	1226	
Db	497	AACNAGCCNAATGGAATAATGAAAAGAGAGGCCACCTTTTNCAGTCAATCTGTCTTC	438	
QY	1227	TC-AAACATAGAAG-AAAGTGCCTATTTTGATT--CCACGTGGAAGGCTTAA-CAGCT	1281	
Db	437	TCNAAACATAGAAGAAAGTGCTATTTTGATTTCACAGTGTGAAGGCTCTAACAGCT	378	
QY	1282	CTAGGAGGCGCTATTTCAGTTTC--TTGTGACCATGTCC-TCACATTGATCAACA	1337	
Db	377	CTAGGAGGCGCTATTTCAGTTTCCTTTGTTGAACATGTCCTCACATTGATCAACA	318	
QY	1338	ATTTAAGATAGAGAGAAAAGAAATCAGAGAAACCTGAAAATGATCATGTGGAGAT	1397	
Db	317	ATTTAAGATAGAGAGAAAAGAAATCAGAGAAACCTGAAAATGATCATGTGGAGAT	258	
QY	1398	TAAGAAGCAGTTGTCC-AAGTATGATCTCAACTTT--CAACAATGAGGAGGAAGTAG	1453	
Db	257	TAGAAGCAGTTGTCCAAAGTATGATCTCAACTTCCACAAAANGAGGAGGAAGTAG	198	
QY	1454	ATACAGATGATCGAACTGAAGGCTATTACAGAGCAGACTCACAGAGCCCTCCCACTTG	1513	
Db	197	ATACAGATGATCGAACTGAAGGCTATTACAGAGCAGACTCACAGAGCCCTCCCACTTG	138	
QY	1514	ATTCTCAGCAGCGCTGCAGTCTTGGAGAGAGAGGTCATGATAGCTCATGCTATCC	1573	
Db	137	ATTCTCAGCAGCGCTGCAGTCTTGGAGAGAGAGGTCATGATAGCTCATGCTATCC	78	
QY	1574	AGGAAGTC-TACAATGATATGTACC	1598	
Db	77	AGGAAGTC-TACAATGATATGTACC	52	

Search completed: September 12, 2003, 22:12:52
Job time : 529.999 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 16:55:50 ; Search time 73 Seconds
(without alignments)
1641.624 Million cell updates/sec

Title: US-09-642-034-5
Perfect score: 4024
Sequence: 1 MARKSLVILITFALSTVNP.....FGIMLLISIFEHKIVRINF 755

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_19Jun03.*			
1:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*		
2:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*		
3:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*		
4:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*		
5:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*		
6:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*		
7:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*		
8:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*		
9:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*		
10:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*		
11:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*		
12:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*		
13:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*		
14:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*		
15:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*		
16:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*		
17:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*		
18:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*		
19:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*		
20:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*		
21:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*		
22:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*		
23:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*		
24:	/SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	4024	100.0	755 23 ABG61889	Prostate cancer-as
2	4024	100.0	755 23 AAM51198	Human breast cancer
3	4024	100.0	755 24 ABR48228	Human bladder cancer
4	4024	100.0	755 24 ABUS6608	Lung cancer-asso
5	4020	99.9	755 23 ABG76949	Human protein, hom
6	4017	99.8	755 22 ABO07206	Human LIV-1-164647
7	4011	99.7	755 23 ABO05550	Breast cancer-asso
8	3969	98.6	749 24 ABR47514	Breast cancer asso
9	3948	98.1	746 23 AAU74631	Oestrogen-regulate

10	3906	97.1	752 24 ABJ37050	Human breast cancer
11	3902	97.0	752 22 AAE07205	Human LIV-1 protei
12	2205	54.8	431 19 AAW34528	Protein encoded by
13	2188	54.4	431 17 AAR98004	Oestrogen-regulate
14	1972	49.0	397 21 AAB56953	Human prostate can
15	1337	33.2	831 22 AAB87345	Human gene 4 encod
16	1337	33.2	831 23 ABG65404	Human albumin fusi
17	1337	33.2	831 23 ABB97347	Novel human-protei
18	1328.5	33.0	829 23 AAU74629	Oestrogen-regulate
19	1153	28.7	298 23 AAM51199	Mouse breast cancer
20	1066	26.5	195 24 ABU70466	Human adipocyte se
21	734	18.2	528 23 AAU74630	Oestrogen-regulate
22	715	17.8	684 22 ABB62691	Drosophila melanog
23	715	17.8	684 23 AAU74625	Oestrogen-regulate
24	657.5	16.3	540 21 AAY57930	Human transmembran
25	652.5	16.2	535 23 AAU74632	Oestrogen-regulate
26	632.5	15.7	492 23 ABG65035	Human albumin fusi
27	632.5	15.7	492 23 AAU91358	Human secreted pro
28	632.5	15.7	492 23 AAU74621	Oestrogen-regulate
29	632.5	15.7	531 20 AAY05376	Human HCMV inducib
30	632.5	15.7	537 22 AAB60496	Oestrogen-regulate
31	632.5	15.7	537 23 AAU74622	Novel human secret
32	607.5	15.1	540 22 AAU30977	Oestrogen-regulate
33	596	14.8	160 22 AAM93460	Human polypeptide,
34	583	14.5	460 22 AAG81272	Human APP protein
35	583	14.5	460 23 ABO74624	Oestrogen-regulate
36	583	14.5	460 23 ABO04723	Human P3105 prote
37	580.5	14.4	462 23 AAU74623	Oestrogen-regulate
38	575.5	14.3	517 23 AAU74627	Oestrogen-regulate
39	575.5	14.3	519 22 ABB71162	Drosophila melanog
40	575.5	14.3	519 23 AAU74628	Oestrogen-regulate
41	573	14.2	393 22 AAB95761	Human protein sequ
42	573	14.2	406 23 ABB89582	Human polypeptide
43	552	13.7	660 23 AAU74620	Oestrogen-regulate
44	552	13.7	691 23 ABP43482	Human secreted pro
45	543.5	13.5	622 23 ABP64821	Human protein seq

ALIGNMENTS

RESULT 1	
ABG61889	
ID	ABG61889 standard; Protein; 755 AA.
XX	ABG61889;
AC	ABG61889;
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	Prostate cancer-associated protein #90.
XX	
KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX	
OS	Mammalia.
XX	
PN	WO200230268-A2.
XX	
PD	18-APR-2002.
XX	
PF	12-OCT-2001; 2001WO-US32045.
XX	
PR	13-OCT-2000; 2000US-0687576.
PR	08-DEC-2000; 2000US-0733288.
PR	08-DEC-2000; 2000US-0733742.
PR	24-JAN-2001; 2001US-263957P.
PR	16-MAR-2001; 2001US-276791P.
PR	16-MAR-2001; 2001US-276889P.
PR	06-APR-2001; 2001US-281922P.
PR	24-APR-2001; 2001US-286214P.
PR	30-APR-2001; 2001US-0847046.
PR	04-MAY-2001; 2001US-288589P.
XX	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.

XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
PI WPI; 2002-471335/50.
XX N-PSDB; ABK92205.
XX Detecting a prostate cancer-associated transcript in a cell in a
PT patient, useful for diagnosing prostate cancer (PC) or screening
PT modulators of PC, by determining if prostate cancer-associated genes
PT are expressed in a prostate tissue -
XX
XX Claim 27; Page 375; 436pp; English.
PS The present invention relates to methods of detecting a prostate
XX cancer-associated transcript in a cell from a patient. The method
XX comprises contacting a biological sample from the patient with
CC prostate cancer-associated polynucleotides (designated PC genes) that
CC selectively hybridize to a sequence that is at least 80% identical
CC to them. The prostate cancer-associated polynucleotide sequences
CC are differentially expressed in prostate tumour tissue or in
CC prostate cancer and are derived from the tissues of various
CC organisms such as humans or other mammals (e.g. mice, sheep and dogs).
CC The methods of the invention are useful for diagnosing and treating
CC prostate cancer in mammals. The prostate cancer-associated genes are
CC useful for diagnosing or treating prostate cancer, as well as for
CC identifying modulators of prostate cancer or agents that inhibit
CC prostate cancer. The nucleic acid sequences are particularly useful
CC in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins.
XX
XX SQ Sequence 755 AA;
Query Match 100.0%; Score 4024; DB 23; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 755; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MARKSVILLITFALSVNPNLHELKAAFPQTTEKISPNWESGINVDIAISTROYHLOOL 60
Db 1 MARKSVILLITFALSVNPNLHELKAAFPQTTEKISPNWESGINVDIAISTROYHLOOL 60
QY 61 FYRGENNSLSVGGPKLLQNGIDKIKRIHIHDDHSDHSDHSDHSDHSDHSDHSDHSDH 120
Db 61 FYRGENNSLSVGGPKLLQNGIDKIKRIHIHDDHSDHSDHSDHSDHSDHSDHSDHSDHSDH 120
QY 121 EHHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDH 180
Db 121 EHHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDHSDH 180
QY 181 SVSASEVTSTVYNTVSEGTHTLETETPRGKLPKDVSSSTPPSVTSKRSVSLAGRKT 240
Db 181 SVSASEVTSTVYNTVSEGTHTLETETPRGKLPKDVSSSTPPSVTSKRSVSLAGRKT 240
QY 241 NESVSEPRKGFMTSRNTNENPOCNASKLITSHGNGIOVPLNATFNPLCPALINQIDA 300
Db 241 NESVSEPRKGFMTSRNTNENPOCNASKLITSHGNGIOVPLNATFNPLCPALINQIDA 300
QY 301 RSLIHTSEKAEIPKTYSLQIAWVGFTAIISTISFLSLGLVILVPLMNRVFPKLLSF 360
Db 301 RSLIHTSEKAEIPKTYSLQIAWVGFTAIISTISFLSLGLVILVPLMNRVFPKLLSF 360
QY 361 LVVALAVGLSGDAFLHLLPHSHASHSHSHSHERPAPMKRGPFLFSLSSQNIIESAYFDS 420
Db 361 LVVALAVGLSGDAFLHLLPHSHASHSHSHSHERPAPMKRGPFLFSLSSQNIIESAYFDS 420
QY 421 TWKGLTALGLYFMFVEHVLTIKQFDKDKKKQKKNKPPENDDDVEIKQLSKYESQLSTN 480
Db 421 TWKGLTALGLYFMFVEHVLTIKQFDKDKKKQKKNKPPENDDDVEIKQLSKYESQLSTN 480
QY 481 EEKVTDORTREGYLRADSOEPHSDSQQPAVLEEEVMTAAHAPQEVNVEYVPRGCKNKC 540
Db 481 EEKVTDORTREGYLRADSOEPHSDSQQPAVLEEEVMTAAHAPQEVNVEYVPRGCKNKC 540
QY 541 HSEFHDITLGQSDDLIHHDHYYHILHHHHQNHHPHSHSQRSYREELKDAGVATLAWMI 600

Db 541 HSEFHDITLGQSDDLIHHDHYYHILHHHHQNHHPHSHSQRSYREELKDAGVATLAWMI 600
QY 601 MGDGLNFDSDGLATGAATFTEGLSSGLSTSVAVFCHPELPHGLDPFVALLKAGMTVKQAVLY 660
Db 601 MGDGLNFDSDGLATGAATFTEGLSSGLSTSVAVFCHPELPHGLDPFVALLKAGMTVKQAVLY 660
QY 661 NALSAMLAIVGATGIFIGHVAENVMWIFALTAGLPMYVALYDMVPEMLHNDASDHGCS 720
Db 661 NALSAMLAIVGATGIFIGHVAENVMWIFALTAGLPMYVALYDMVPEMLHNDASDHGCS 720
QY 721 RWGYFFLQAGMLGFGIMLLISIFEKIVFRINF 755
Db 721 RWGYFFLQAGMLGFGIMLLISIFEKIVFRINF 755
RESULT 2
AAM51198
ID AAM51198 standard; Protein: 755 AA.
XX AAM51198;
XX AC AAM51198;
XX XX
DT 10-JUN-2002 (first entry)
XX Human breast cancer 4 gene (BCR4)-encoded protein.
DE BCR4; human; breast cancer 4 gene; prostate cancer; diagnosis;
KW therapy; vaccine.
XX Homo sapiens.
XX XX
FH Key
FT Peptide 1..20
FT Protein 21..755
FT Domain 330..346
FT Domain 352..368
FT Domain 427..444
FT Domain 663..679
FT Domain 688..703
FT Domain 730..745
FT Domain 730..745
XX WO200216939-A2.
XX 28-FEB-2002.
XX 20-AUG-2001; 2001WO-US25997.
XX 18-AUG-2000; 2000US-0642034.
XX 08-DEC-2000; 2000US-0733320.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Mack D, Gish KC, Wilson KE;
XX WPI; 2002-242033/29.
XX N-PSDB; ABA92299.
XX Screening for drugs that affect expression of the breast cancer 4 gene
PT or its fragments, use of these to treat prostate and breast cancer, and
PT diagnosing these diseases -
XX Example 4; Fig 2; 83pp; English.
XX The presence sequence is that of the protein encoded by the human
CC breast cancer 4 gene (BCR4) on chromosome 18 (see ABA92299). The

QY 421 TWKGLTALGGLYPMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKQSKYESQISTN 480
 DB 421 TWKGLTALGGLYPMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKQSKYESQISTN 480
 QY 481 EEKVDTDRTEGYLRADSOEFSHDSQPAVLEEEVMIHAHPQEVNVEYVPRGCKNC 540
 DB 481 EEKVDTDRTEGYLRADSOEFSHDSQPAVLEEEVMIHAHPQEVNVEYVPRGCKNC 540
 QY 541 HSHFHDPLGQSDLLIHHDYHHILHHHQNHPHSHSQRSREELKDAGVATLAWVI 600
 DB 541 HSHFHDPLGQSDLLIHHDYHHILHHHQNHPHSHSQRSREELKDAGVATLAWVI 600
 QY 601 MGDGLHNFSDGLAIGAAFTGSLSTSVAVFCHPELGHDFAVLLKAGMTVKQAVLY 660
 DB 601 MGDGLHNFSDGLAIGAAFTGSLSTSVAVFCHPELGHDFAVLLKAGMTVKQAVLY 660
 QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS 720
 DB 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS 720
 QY 721 RWGYFFLQAGMLLGFIMLLISIFEHKIVRINF 755
 DB 721 RWGYFFLQAGMLLGFIMLLISIFEHKIVRINF 755
 RESULT 6
 AA07206
 ID AAE07206 standard; Protein; 755 AA.
 XX AC AA07206;
 XX DT 06-NOV-2001 (first entry)
 XX DE Human LIV-1-164647 protein.
 XX KW Human; LIV-1-164647; cytostatic; estrogen-inducible gene; tumour;
 KW cancer; breast; lung; prostate; colon; ovary; uterus; kidney; gastric;
 KW salivary gland; carcinoma; drug screening; therapy.
 XX OS Homo sapiens.
 XX PN W0200155178-A2.
 XX PD 02-AUG-2001.
 XX PF 25-JAN-2001; 2001WO-US02622.
 XX PR 25-JAN-2000; 2000US-0177951.
 XX PR 10-APR-2000; 2000US-0195761.
 XX FA (GETH) GENENTECH INC.
 XX PI Goddard A, Gurney AL, Smith V, Hongo JS, De Sauvage F;
 XX WPI; 2001-502628/55.
 XX DR N-PSDB; AAD13480.
 XX PT New LIV-1 DNA and amino acids, useful for treating cancer or inhibiting
 PT the proliferation of tumor cells in mammals, e.g. breast, lung,
 PT prostate, colon, ovary, uterus, kidney, gastric or salivary carcinoma
 PT .
 XX Claim 20; Fig 2B; 150pp; English.
 XX CC The present sequence is human estrogen-inducible LIV-1-164647
 CC protein. LIV-1 is overexpressed in tumour tissues such as prostate,
 CC colon, lung and breast. The LIV-1 DNA and polypeptide, and the monoclonal
 CC antibody are useful for treating cancer and inhibiting the proliferation
 CC of tumour cells in mammals, e.g. breast, lung, prostate, colon, ovary,
 CC uterus, kidney, gastric or salivary carcinoma, or other tumour cell
 CC types expressing the LIV-1-164647 protein. In particular, the mammal is
 CC a human. The LIV-1 DNA and polypeptide may also be used in screening
 CC assays for drug candidates.

XX SQ Sequence 755 AA;
 Query Match 99.8%; Score 4017; DB 22; Length 755;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 754; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MARKLSVILLIFALSVTNPLHKLKAAAPQTEKISPNWESGINVDLAISTRQYHLQOL 60
 DB 1 MARKLSVILLIFALSVTNPLHKLKAAAPQTEKISPNWESGINVDLAISTRQYHLQOL 60
 QY 61 FYRYGNNLSVEGRKLLQNGIDIKRIHHHDHSDHEHSHDSHDSHDSHDSHDSHDSH 120
 DB 61 FYRYGNNLSVEGRKLLQNGIDIKRIHHHDHSDHEHSHDSHDSHDSHDSHDSHDSH 120
 QY 121 EHSDDHDSHSHNHAASGKRRKALCPDHDSDSSSKDPNSQSGAHRPEHAGRRNVKD 180
 DB 121 EHSDDHDSHSHNHAASGKRRKALCPDHDSDSSSKDPNSQSGAHRPEHAGRRNVKD 180
 QY 181 SVSASEVTSTVNTVSEGTETETETETETETETETETETETETETETETETETETET 240
 DB 181 SVSASEVTSTVNTVSEGTETETETETETETETETETETETETETETETETETETET 240
 QY 241 NESYSEPRKGFMYSRNTNENQECFNASKLLTSHGMIQVPLNATEFNYLCPALINQIDA 300
 DB 241 NESYSEPRKGFMYSRNTNENQECFNASKLLTSHGMIQVPLNATEFNYLCPALINQIDA 300
 QY 301 RSLIHTSEKKAIEPPKTYSLQIAWVGGFIAISLISFLSLLGVLLVPLMNRVFFKLLSF 360
 DB 301 RSLIHTSEKKAIEPPKTYSLQIAWVGGFIAISLISFLSLLGVLLVPLMNRVFFKLLSF 360
 QY 361 LVAVAGTILSGDAFLHLLPHSHASHSHSHSHEPAMENKRGPLFSLSSONIEESAYFDS 420
 DB 361 LVAVAGTILSGDAFLHLLPHSHASHSHSHSHEPAMENKRGPLFSLSSONIEESAYFDS 420
 QY 421 TWKGLTALGGLYPMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKQSKYESQISTN 480
 DB 421 TWKGLTALGGLYPMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKQSKYESQISTN 480
 QY 481 EEKVDTDRTEGYLRADSOEFSHDSQPAVLEEEVMIHAHPQEVNVEYVPRGCKNC 540
 DB 481 EEKVDTDRTEGYLRADSOEFSHDSQPAVLEEEVMIHAHPQEVNVEYVPRGCKNC 540
 QY 541 HSHFHDPLGQSDLLIHHDYHHILHHHQNHPHSHSQRSREELKDAGVATLAWVI 600
 DB 541 HSHFHDPLGQSDLLIHHDYHHILHHHQNHPHSHSQRSREELKDAGVATLAWVI 600
 QY 601 MGDGLHNFSDGLAIGAAFTGSLSTSVAVFCHPELGHDFAVLLKAGMTVKQAVLY 660
 DB 601 MGDGLHNFSDGLAIGAAFTGSLSTSVAVFCHPELGHDFAVLLKAGMTVKQAVLY 660
 QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS 720
 DB 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS 720
 QY 721 RWGYFFLQAGMLLGFIMLLISIFEHKIVRINF 755
 DB 721 RWGYFFLQAGMLLGFIMLLISIFEHKIVRINF 755
 RESULT 7
 ABJ05550
 ID ABJ05550 standard; Protein; 755 AA.
 XX AC ABJ05550;
 XX DT 14-NOV-2002 (first entry)
 XX DE Breast cancer-associated protein 15.
 XX KW Breast cancer; breast cancer-associated gene sequence;
 KW drug development; pharmacogenetics; biosensor development.
 XX

QY	121	EHHSDHDDHHSHHNAASGKNKKKALCPDHDSDSGKDPNRSQKGAHPEHASGRRNVKD	180
Db	121	EHHSDD-----HNAASGKNKKKALCPDHDSDSGKDPNRSQKGAHPEHASGRRNVKD	174
QY	181	SVSASEVTSVYNTVSEGTHTLETETPRCKLFPKDVSSSTPPSVTSKSRVSLAGRKT	240
Db	175	SVSASEVTSVYNTVSEGTHTLETETPRCKLFPKDVSSSTPPSVTSKSRVSLAGRKT	234
QY	241	NESVSEPRKGFMTSRNTNENPOCFNASKLLTSHGMGIQVPLNATFNWYLCPALINQIDA	300
Db	235	NESVSEPRKGFMTSRNTNENPOCFNASKLLTSHGMGIQVPLNATFNWYLCPALINQIDA	294
QY	301	RSLIHTSEKKABIPKTYSLQIAWVGGFIAISIIISFLSILGLVILVPLMNRVFKFLLSF	360
Db	295	RSLIHTSEKKABIPKTYSLQIAWVGGFIAISIIISFLSILGLVILVPLMNRVFKFLLSF	354
QY	361	LYALAVGTILSGDAFLHLLPHSHASHHSHSEHPAMEMKRGPLFSLHSSQNIBESAYFDS	420
Db	355	LYALAVGTILSGDAFLHLLPHSHASHHSHSEHPAMEMKRGPLFSLHSSQNIBESAYFDS	414
QY	421	TWKGLTALGLIYPMFJUEVJFLIKQFKDKKKKNQKKPENDDDVEIKKQLSKYESQLSTN	480
Db	415	TWKGLTALGLIYPMFJUEVJFLIKQFKDKKKKNQKKPENDDDVEIKKQLSKYESQLSTN	474
QY	481	EKKYDTRDTREGYLRADSQEPSHFDSSQPAVLEEVEVMTAHAPQEVYNETVPRGCKNK	540
Db	475	EKKYDTRDTREGYLRADSQEPSHFDSSQPAVLEEVEVMTAHAPQEVYNETVPRGCKNK	534
QY	541	HSHPDHTLGSDDLIIHHHDYHILIIHHHQNHHPHSHSQRYSRBELKDAGVATLAWVI	600
Db	535	HSHPDHTLGSDDLIIHHHDYHILIIHHHQNHHPHSHSQRYSRBELKDAGVATLAWVI	594
QY	601	MGDGLHNFSDGLAIGAFTGSLGSTSVAVFCHLPHLGDFAVLILKAGMTVKQAVLY	660
Db	595	MGDGLHNFSDGLAIGAFTGSLGSTSVAVFCHLPHLGDFAVLILKAGMTVKQAVLY	654
QY	661	NALSAMLAYLGMATGIFTIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS	720
Db	655	NALSAMLAYLGMATGIFTIGHYAENVSMWIFALTAGLFMTVALVDMVPEMLHNDASDHGCS	714
QY	721	RWGYFFFLQNAQMLLGGIMLLISIFEHKIYF	751
Db	715	RWGYFFFLQNAQMLLGGIMLLISIFEHKIYF	745
RESULT 10			
ABJ37050			
ID	ABJ37050 standard; Protein; 752 AA.		
AC	XX		
AD	ABJ37050;		
DT	01-MAY-2003 (first entry)		
XX	XX		
DE	Human breast cancer / ovarian cancer related protein #26.		
XX	XX		
KW	Human; cytostatic; breast cancer; ovarian cancer.		
OS	Homo sapiens.		
XX	XX		
PN	WO2003000012-A2.		
XX	XX		
PD	03-JAN-2003.		
XX	XX		
PF	21-JUN-2002; 2002WO-US19773.		
XX	XX		
PR	21-JUN-2001; 2001US-300159P.		
XX	XX		
PR	27-JUN-2001; 2001US-301351P.		
XX	XX		
PA	(MILL-) MILLENNIUM PHARM INC.		
XX	XX		
PI	Veibj OP;		
XX	XX		
DR	WPI; 2003-267848/26.		

```

RESULT 11
AAE07205
ID AAE07205 standard; Protein; 752 AA.
XX
AC AAE07205;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human LIV-1 protein.
XX
KW Human; LIV-1; cytostatic; estrogen-inducible gene; tumour; therapy;
KW cancer; breast; lung; prostate; colon; ovary; uterus; kidney; gastric;
KW salivary gland; carcinoma; drug screening.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= Signal_peptide
FT Protein 21..752
FT /label= Mature_LIV-1_protein
FT Domain 24..312
FT /label= Extracellular_domain
FT Domain 318..367
FT /label= Transmembrane_domain
XX
WC200155178-A2.
XX
XX 02-AUG-2001.
XX
XX 25-JAN-2001; 2001WO-US02622.
XX
XX 25-JAN-2000; 2000US-0177951.
XX 10-APR-2000; 2000US-0195761.
XX
XX (SETH ) GENENTECH INC.
XX
XX Goddard A, Gurney AL, Smith V, Hongo JS, De Sauvage F;
XX
XX WPI: 2001-502628/55.
XX N-PSDB; AAD13479.
XX
XX New LIV-1 DNA and amino acids, useful for treating cancer or inhibiting
XX the proliferation of tumor cells in mammals, e.g. breast, lung,
XX prostate, colon, ovary, uterus, kidney, gastric or salivary carcinoma
XX
XX Example 2; Fig 1B; 150pp; English.
XX
XX The present sequence is human estrogen-inducible protein (LIV-1).
XX LIV-1 is overexpressed in tumour tissues such as prostate, colon,
XX lung and breast. The LIV-1 DNA and polypeptide, and the monoclonal
XX antibody are useful for treating cancer and inhibiting the proliferation
XX of tumour cells in mammals, e.g. breast, lung, prostate, colon, ovary,
XX uterus, kidney, gastric or salivary carcinoma, or other tumour cell
XX types expressing the LIV-1-164647 protein. In particular, the mammal is
XX a human. The LIV-1 DNA and polypeptide may also be used in screening
XX assays for drug candidates.
XX
XX Sequence 752 AA;
XX
Query Match 97.0%; Score 3902; DB 22; Length 752;
Best Local Similarity 99.2%; Pred. NO. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
XX
QY 1 MARKLSVLIITLFTALSVNPLHELKAAAFPTTEKISPNWESGINVDLAISTROYHLQQL 60
DB 1 MARKLSVLIITLFTALSVNPLHELKAAAFPTTEKISPNWESGINVDLAISTROYHLQQL 60
XX
QY 61 FYRGNNLSLVEGFRKLQNLGIDKIKRIHHDDHSDHSDHEHSDHEHSDH 120
DB 61 FYRGNNLSLVEGFRKLQNLGIDKIKRIHHDDHSDHSDHEHSDHEHSDH 120

```

```

QY 121 EHHSDHSHSHENHAASGKNRKALCPDHDSDSSGKDPNSQKGARHPEHAGSRNNVKD 180
DB 121 EHHSD-----HNHAASGKNRKALCPDHDSDSSGKDPNSQKGARHPEHAGSRNNVKD 174
QY 181 SVSASEVTSTVNTVSGTHFLETIETPRPGKLPFKDVSSSTPSPVTSKSRVSLAGRKT 240
DB 175 SVSASEVTSTVNTVSGTHFLETIETPRPGKLPFKDVSSSTPSPVTSKSRVSLAGRKT 234
QY 241 NESVSEPRKGFMYSRNTNENPQCFNASKLILSHGMIQVPLNATEFNILCPALINQIDA 300
DB 235 NESVSEPRKGFMYSRNTNENPQCFNASKLILSHGMIQVPLNATEFNILCPALINQIDA 294
QY 301 RSLIITSEKKAIEPPKTYISLQIAWGGPTAISISFLSILGLVILVPLMRVFFKLLSF 360
DB 295 RSLIITSEKKAIEPPKTYISLQIAWGGPTAISISFLSILGLVILVPLMRVFFKLLSF 354
QY 361 LVALAVGTLTSGDAFLHLPLPHSHASHHSHBEPAMEMKRGFLPSHLSSONIEESAYFDS 420
DB 355 LVALAVGTLTSGDAFLHLPLPHSHASHHSHBEPAMEMKRGFLPSHLSSONIEESAYFDS 414
QY 421 TWKGLTALGGLYFMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKOLSKYESQLSTN 480
DB 415 TWKGLTALGGLYFMFLVEHVLTLIKQFKDKKKKQKPPENDDDVEIKKOLSKYESQLSTN 474
QY 481 EEKVDTDRTGEGYLRADSOEPEHFDSSQPAVLEEEVMAIAHAHPQEVYNEYVPRGCKNC 540
DB 475 EEKVDTDRTGEGYLRADSOEPEHFDSSQPAVLEEEVMAIAHAHPQEVYNEYVPRGCKNC 534
QY 541 HSHFHTLQGSDDLHHHHHHYHLLHHHHHNNHHPHSHSQRYSREELKDAGVATLAWWVI 600
DB 535 HSHFHTLQGSDDLHHHHHHYHLLHHHHHNNHHPHSHSQRYSREELKDAGVATLAWWVI 594
QY 601 MGDGLNFDGLAIGAAPTGLSSGLSTSVAVFCHLPHLGLDFAVLLKAGMTVKQAVLY 660
DB 595 MGDGLNFDGLAIGAAPTGLSSGLSTSVAVFCHLPHLGLDFAVLLKAGMTVKQAVLY 654
QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALVDMVPEMLHNDASDHGCS 720
DB 655 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALVDMVPEMLHNDASDHGCS 714
QY 721 RWGYFFLQAGMLLGFGLMILLI 742
DB 715 RWGYFFLQAGMLLGFGLMILLI 736
XX
RESULT 12
AAW34528
ID AAW34528 standard; Protein; 431 AA.
XX
AC AAW34528;
XX
DT 24-MAR-1998 (first entry)
XX
DE Protein encoded by pLIV1 gene partial sequence.
XX
KW Oestrogen regulated gene; pLIV1 gene partial sequence; breast cancer;
KW metastatic risk assessment; hormone therapy responsiveness;
KW tumour oestrogen receptor expression; lymph node involvement.
XX
XX Homo sapiens.
XX
XX US5693465-A.
XX
XX 02-DEC-1997.
XX
XX 22-SEP-1994; 94US-0311023.
XX
XX 22-SEP-1994; 94US-0311023.
XX
XX (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
XX
XX Gee JMW, Green CD, Manning DL, Nicholson RI;
XX

```

DR WPI; 1998-031740/03.
 DR N-PSDB; AAT99070, AAT99071.
 XX Assessment of metastatic risk or oestrogen responsive-ness in breast
 FT cancer - by detecting expression of pLIV1 protein
 XX
 PS Claim 1; Column 5-10; 11pp; English.
 XX
 CC This sequence is encoded by a partial sequence of the pLIV1 gene. This
 CC protein sequence is an oestrogen regulated gene associated with breast
 CC cancer. This sequence can be used in the method of the invention. The
 CC method is for determining the risk of metastasis of a female breast
 CC tumour and predicting the responsiveness of a female breast
 CC hormone treatment comprise determining if a tissue sample from the tumour
 CC expresses a polypeptide comprising at least 14 continuous amino acids of
 CC the protein encoded by this sequence. The method is useful for
 CC determining the risk of metastasis of a female breast tumour and to
 CC predict the responsive of the female breast tumour to hormone treatment.
 CC pLIV1 expression is highly predictive of tumour oestrogen receptor
 CC expression and lymph node involvement.
 XX
 XX Sequence 431 AA;
 SQ
 Query Match 54.8%; Score 2205; DB 19; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.2e-198;
 Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 328 GFTAIISIISFLSLGLVILPLMNRVFFKLLSFLVALAVGTLSDAFLLHLLPHSHASHH 387
 Db 1 GFTAIISIISFLSLGLVILPLMNRVFFKLLSFLVALAVGTLSDAFLLHLLPHSHASHH 60
 QY 388 SHSHEEPAMKRGFLPSHLSNIEESAYFDSTWKGTLALGGLYFMFLVEHVLIILKQF 447
 Db 61 SHSHEEPAMKRGFLPSHLSNIEESAYFDSTWKGTLALGGLYFMFLVEHVLIILKQF 120
 QY 448 KDKKKKNOKKPENNDDVEIKKQSKYESQSTNEEKVDTDDRTGTYLRADSOEPSHFDSQ 507
 Db 121 KDKKKKNOKKPENNDDVEIKKQSKYESQSTNEEKVDTDDRTGTYLRADSOEPSHFDSQ 180
 QY 508 QPAVLEEEVMIATAHAHPQEVYNEYVPRGCKNCKCHSHFDHILGQSDLLIHHDYHILH 567
 Db 181 QPAVLEEEVMIATAHAHPQEVYNEYVPRGCKNCKCHSHFDHILGQSDLLIHHDYHILH 240
 QY 568 HHQNHHPHSHSQSYREELKDAGVATLAWMYINGDGLHNFSDGLAIGRAFTGLSSGLS 627
 Db 241 HHQNHHPHSHSQSYREELKDAGVATLAWMYINGDGLHNFSDGLAIGRAFTGLSSGLS 300
 QY 628 TSVAVFCHELPHLGDFAVLLKAGMTVKQAVLYNALSAMLAYLGMATGIFIGHYAENVSM 687
 Db 301 TSVAVFCHELPHLGDFAVLLKAGMTVKQAVLYNALSAMLAYLGMATGIFIGHYAENVSM 360
 QY 688 WIFALTAGLFWYVALDMPVEMLHNDASDHGCSRWGFFFLQAGMLLIGRGIMILLI 742
 Db 361 WIFALTAGLFWYVALDMPVEMLHNDASDHGCSRWGFFFLQAGMLLIGRGIMILLI 415
 RESULT 13
 AAR98004
 ID AAR98004 standard; Protein; 431 AA.
 XX AC
 XX AAR98004;
 XX
 DT 16-OCT-1996 (first entry)
 XX
 DE Oestrogen-regulated pLIV1 gene-encoded protein.
 XX
 KW Oestrogen-regulated gene; pLIV1; breast cancer; tumour; diagnosis;
 KW marker; metastasis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 125..138

FT /label=1003
 FT /note="immunogenic peptide 1003, claim 11,
 FT page 19"
 FT Peptide 252..265
 FT /label=1004
 FT /note="immunogenic peptide 1004, claim 11
 FT page 19"
 FT Peptide 418..431
 FT /label=1105
 FT /note="immunogenic peptide 1005, claim 11
 FT page 19"
 FT
 PN CA2132500-A.
 XX
 PD 21-MAR-1996.
 XX
 PF 20-SEP-1994; 94CA-2132500.
 XX
 PR 20-SEP-1994; 94CA-2132500.
 XX
 PA (UYWA-) UNIV WALES COLLEGE OF MEDICINE.
 XX
 PI Gee JM, Green CD, Manning DL, Nicholson RI;
 XX
 DR WPI; 1996-268987/28.
 DR N-PSDB; AAT33219, AAT33220.
 XX
 FT Oestrogen-regulated pLIV1 gene and deduced polypeptide - useful for
 FT predicting the propensity for metastatic spread or the
 FT responsiveness to endocrine treatment of breast tumour
 XX
 PS Claim 1; Page 14-16; 27pp; English.
 XX
 CC The polypeptide (AAR98004) encoded by a partial cDNA clone
 CC (AAT33219) corresponding to the human pLIV1 gene contains 3 peptide
 CC regions, 1003, 1004 and 1005, that were used to raise polyclonal
 CC antibodies in rabbits. Peptides 1004 and 1005 were most
 CC immunogenic, generating ELISA-positive sera at dilutions of
 CC 1/20000. Cross-reactivity studies using the 3 peptides identified
 CC 7 positive specific sera in tumour samples from 74 patients with
 CC primary breast cancer. The pLIV1 polypeptide was expressed in the
 CC cytosol of oestrogen receptor-positive epithelial cells of breast
 CC tumours. Its presence can be used to predict the propensity for
 CC metastatic spread or the responsiveness of a breast tumour to
 CC endocrine treatment. Its inactivation may be used as a means of
 CC treating breast cancer and metastasis.
 XX
 SQ Sequence 431 AA;
 Query Match 54.4%; Score 2188; DB 17; Length 431;
 Best Local Similarity 99.3%; Pred. No. 4.8e-197;
 Matches 412; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 328 GFTAIISIISFLSLGLVILPLMNRVFFKLLSFLVALAVGTLSDAFLLHLLPHSHASHH 387
 Db 1 GFTAIISIISFLSLGLVILPLMNRVFFKLLSFLVALAVGTLSDAFLLHLLPHSHASHH 60
 QY 388 SHSHEEPAMKRGFLPSHLSNIEESAYFDSTWKGTLALGGLYFMFLVEHVLIILKQF 447
 Db 61 SHSHEEPAMKRGFLPSHLSNIEESAYFDSTWKGTLALGGLYFMFLVEHVLIILKQF 120
 QY 448 KDKKKKNOKKPENNDDVEIKKQSKYESQSTNEEKVDTDDRTGTYLRADSOEPSHFDSQ 507
 Db 121 KDKKKKNOKKPENNDDVEIKKQSKYESQSTNEEKVDTDDRTGTYLRADSOEPSHFDSQ 180
 QY 508 QPAVLEEEVMIATAHAHPQEVYNEYVPRGCKNCKCHSHFDHILGQSDLLIHHDYHILH 567
 Db 181 QPAVLEEEVMIATAHAHPQEVYNEYVPRGCKNCKCHSHFDHILGQSDLLIHHDYHILH 240
 QY 568 HHQNHHPHSHSQSYREELKDAGVATLAWMYINGDGLHNFSDGLAIGRAFTGLSSGLS 627
 Db 241 HHQNHHPHSHSQSYREELKDAGVATLAWMYINGDGLHNFSDGLAIGRAFTGLSSGLS 300

QY 628 TSVAVFCHPELHDFVALLKAGMTVKQAVLYNALNSAMLAYLGMATGIFIGHYAENVSM 687
 DB 301 TSVAVFCHPELHDFVALLKAGMTVKQAVLYNALNSAMLAYLGMATGIFIGHYAENVSM 360
 QY 688 WIFALFAGLFMTVALVDMVPEMLHNDASDHGCSRWGYFFLQWAGMLLGFGLMILLI 742
 DB 361 WIFALFAGLFMTVALVDMVPEMLHNDASDHGCSRWGYFFLQWAGMLLGFGLMILLI 415

RESULT 14
 AAB56953
 ID AAB56953 standard; Protein; 397 AA.
 XX AC AAB56953;
 XX 13-MAR-2001 (first entry)
 XX Human prostate cancer antigen protein sequence SEQ ID NO:1531.
 DE Human prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytostatic; cardioactive; immunomodulatory; muscular;
 KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;
 KW antibacterial; gene therapy; neural; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.
 XX Homo sapiens.
 OS
 XX WO2000055174-A1.
 PN 21-SEP-2000.
 PD 08-MAR-2000; 2000WO-US05988.
 PF 12-MAR-1999; 99US-0124270.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM;
 XX WPI; 2000-587513/55.
 DR N-PSDB; AAF16156.
 XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX Claim 11; Page 1974-1975; 2338pp; English.
 XX AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytostatic,
 CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
 CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 XX Sequence 397 AA;
 SQ Query Match 49.0%; Score 1972; DB 21; Length 397;
 Best Local Similarity 98.7%; Pred. No. 9.5e-177;
 Matches 376; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 161 SVSASEVTSTVNTYSEGTHLETETPRPGKLFPKDVSSSTPPSVTSKSRVSLAGRKT 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 SVSASEVTSTVNTYSEGTHLETETPRPGKLFPKDVSSSTPPSVTSKSRVSLAGRKT 60
 QY 241 NESVSPRKGFMYSNTNENPQECFNASKLLTSHGNGIQVPLNATERNYLCPALIIQIDA 300
 DB 61 NESVSPRKGFMYSNTNENPQECFNASKLLTSHGNGIQVPLNATERNYLCPALIIQIDA 120
 QY 301 RSLIHTSSKKAIEPPKTYSLQIAWVGFTASISFSLGLLVILPLMNRVFFKLLSF 360
 DB 121 RSLIHTSSKKAIEPPKTYSLQIAWVGFTASISFSLGLLVILPLMNRVFFKLLSX 180
 QY 361 LVALAVGTLTSGDAFLHLLPHSHASHSHSHEEPAMEMKRGFLFSLSSQNIIESAYFDS 420
 DB 181 XVALAVGTLTSGDAFLHLLPHSHASHSHSHEEPAMEMKRGFLFSLSSQNIIESAYFDS 240
 QY 421 TWKGLTALGGLYFMFLVEHVLTLIKQFKDKKKKNKPPENDDDVEIKKQSKYESQLSTN 480
 DB 241 TWKGLTALGGLYFMFLVEHVLTLIKQFKDKKKKNKPPENDDDVEIKKQSKYESQLSTN 300
 QY 481 EEKVYDTRDTEGYLRADSOEPSHFDSQOPAVLEEEVMTIAHAHQEVNYEYVPRGCKNC 540
 DB 301 EEKVYDTRDTEGYLRADSOEPSHFDSQOPAVLEEEVMTIAHAHQEVNYEYVPRGCKKC 360
 QY 541 HSHFHTLQGSDDLHIIHHHDY 561
 DB 361 HSHFHTLQGSDDLHIIHHDF 381

RESULT 15
 AAB87345
 ID AAB87345 standard; Protein; 831 AA.
 XX AC AAB87345;
 XX 22-MAY-2001 (first entry)
 XX Human gene 4 encoded secreted protein HDPFY41, SEQ ID NO:86.
 DE Human; secreted protein; proliferative disorder; cancer; tumour;
 KW foetal abnormality; developmental abnormality; haematopoietic disorder;
 KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiogenic disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder; infection; wound healing; vulnary;
 KW cell culture; chemotaxis; food additive;
 KW binding partner identification.
 XX Homo sapiens.
 OS
 XX WO200118022-A1.
 PN 15-MAR-2001.
 PD 31-AUG-2000; 2000WO-US24008.
 XX 03-SEP-1999; 99US-0152315.
 PR 03-SEP-1999; 99US-0152317.
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
 PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
 PI Moore PA, Shi Y, Wei Y, Florence KA;
 XX WPI; 2001-203081/20.
 DR N-PSDB; AAF91861.
 XX Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 XX

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 22:13:01 ; Search time 40 seconds
(without alignments)
1815.184 Million cell updates/sec

Title: US-09-642-034-5
Perfect score: 4024
Sequence: 1 MARKLSVILITFALSVTNP.....FGIMLLISIFEHKIVIRINE 755

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3902	97.0	752	G02273	LIV-1 protein - hu
2	316	15.3	529	T08684	hypothetical prote
3	393.5	9.8	513	T23089	hypothetical prote
4	391	9.7	360	T19585	hypothetical prote
5	381.5	9.5	586	T24835	hypothetical prote
6	376	9.3	404	T25420	hypothetical prote
7	325	8.1	436	I49714	MHC H-2K/t-w5-link
8	299	7.4	346	S49599	probable membrane
9	271.5	6.7	433	T39240	hypothetical prote
10	250.5	6.2	450	C96704	unknown protein, 2
11	208	5.2	338	B89625	protein C14H10.1 f
12	208	5.2	362	T19285	hypothetical prote
13	194.5	4.8	735	T45059	hypothetical prote
14	180	4.5	3119	T18414	protein g377 - mal
15	178	4.4	144	C44863	hypothetical prote
16	171.5	4.3	668	A44863	trophozoite antige
17	168.5	4.2	764	H71607	trophozoite antige
18	168	4.2	351	KGQHL	histidine-rich gly
19	166.5	4.1	269	H71107	gufa protein homol
20	163	4.1	65	D44863	trophozoite antige
21	163	4.1	283	C85838	hypothetical prote
22	161.5	4.0	972	S35521	DNA topoisomerase
23	160.5	4.0	477	T29592	hypothetical prote
24	158	3.9	140	A54523	histidine-rich pro
25	157	3.9	257	AH1865	hypothetical prote
26	155	3.9	279	H90992	hypothetical prote
27	155	3.9	295	T15544	hypothetical prote
28	155	3.9	330	D90548	hypothetical prote
29	154.5	3.8	96	A44971	hypothetical prote

30 154.5 3.8 1225 2 T16346
31 154.5 3.8 2910 2 T28156
32 154 3.8 337 2 D64049
33 152 3.8 259 1 A69162
34 147.5 3.7 291 2 F81444
35 146.5 3.6 905 2 AC2680
36 146.5 3.6 916 2 A97462
37 146 3.6 427 2 A32372
38 146 3.6 549 2 B32372
39 146 3.6 1510 2 T33100
40 146 3.6 1785 2 T21558
41 146 3.6 1929 2 T21559
42 145.5 3.6 348 2 T06385
43 145 3.6 254 1 S39876
44 143.5 3.6 355 2 T52183
45 143 3.6 327 2 S31781

ALIGNMENTS

RESULT 1
G02273
LIV-1 protein - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 05-Nov-1999
C:Accession: G02273
R:Green, C.; Gilhooly, E.M.; Walker, N.J.
submitted to the EMBL Data Library, November 1995
A:Reference number: H00961
A:Accession: G02273
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-752 <GRE>
A:Cross-references: EMBL:U41060; NID:g1256000; PIDN:AAA96258.1; PID:g1256001
C:Genetics:
A:Gene: LIV-1

Query Match 97.0%; Score 3902; DB 2; Length 752;
Best Local Similarity 99.2%; Pred. No. 2.6e-277;
Matches 736; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 1 MARKLSVILITFALSVTNPHELKAAAFPTTEKISPNWESGINVDLAISTROYHLQOL 60
Db 1 MARKLSVILITFALSVTNPHELKAAAFPTTEKISPNWESGINVDLAISTROYHLQOL 60
Qy 61 FYRYGENNSLVEGFRKLLQNIQIDKIKRIHHDDHSDHEHSDHEHSDHEHSDH 120
Db 61 FYRYGENNSLVEGFRKLLQNIQIDKIKRIHHDDHSDHEHSDHEHSDHEHSDH 120
Qy 121 EHSDDHSHHNAAGSKNKKALCPDHDSDSSGKDPNRSQKGARPEHASGRNVKD 180
Db 121 EHSDD-----HNHAASGKNKKALCPDHDSDSSGKDPNRSQKGARPEHASGRNVKD 174
Qy 181 SVSASEVTSTVYNVSEGGHFELETETPRPKLPKDVSSSTPPSTYSKRSYRLAGKKT 240
Db 175 SVSASEVTSTVYNVSEGGHFELETETPRPKLPKDVSSSTPPSTYSKRSYRLAGKKT 234
Qy 241 NESVSEPRKGFMYSRNTNENPOECFNASKLLTSHGMGIQVPLNATFNFLCPAILNQIDA 300
Db 235 NESVSEPRKGFMYSRNTNENPOECFNASKLLTSHGMGIQVPLNATFNFLCPAILNQIDA 294
Qy 301 RSLIHSSEKKAETPKTYSLQIAVWGFFIAISLISFLSLGLVILVPLMNRVFFKLLSF 360
Db 295 RSLIHSSEKKAETPKTYSLQIAVWGFFIAISLISFLSLGLVILVPLMNRVFFKLLSF 354
Qy 361 LVVALAVGTLSGDAFLHLLPHSHASHHHSHSEEPAMEMKRGPLFSLHSQNIIESAYFDS 420
Db 355 LVVALAVGTLSGDAFLHLLPHSHASHHHSHSEEPAMEMKRGPLFSLHSQNIIESAYFDS 414
Qy 421 TWKGTALGGLYFMFLVHVLTLIKQFKDKKKKKKKPENNDDVDVEIKKOLSKYESQLSTN 480
Db 415 TWKGTALGGLYFMFLVHVLTLIKQFKDKKKKKKKPENNDDVDVEIKKOLSKYESQLSTN 474

Db 388 ASFIAGTGVIGTMITVTVLHVPEHIGDFAILIQSGYSKKKAMLIQLVTA-----LGALSG 443
QY 576 IFIGHYAE-----VSMWIFALTAGLFWTVALVDMPEMLHNDASDHGSGRWGYFFL 727
Db 444 CVISLFSADADALAAASWVLPFTAGGFIYIATVSVIPELLENGS-----FF 492
QY 728 QNA-----GMLIGFGIMLLISIEF 746
Db 493 QTVKEIFAILTGIFIMLYLAIYE 515

RESULT 4
T19585
hypothetical protein C30H6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
A:Accession: T19585
R:Mortimore, B.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19148
A:Accession: T19585
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-360 <WIL>
A:Cross-references: EMBL:Z81044; PIDN:CAB02806.1; GSPDB:GN00022; CESP:C30H6.2
A:Experimental source: clone C30H6
C:Genetics:
A:Gene: CESP:C30H6.2
A:Map position: 4
A:Introns: 43/2; 67/3; 111/3; 149/3; 231/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C30H6.2

Query Match 9.7%; Score 391; DB 2; Length 360;
Best Local Similarity 28.0%; Pred. No. 6.3e-21;
Matches 112; Conservative 56; Mismatches 136; Indels 94; Gaps 11;
QY 328 GFAISISFSLGLVILPLMNRVFKFLSLFVALAVGLSDAPLHLLP-----HSH 382
Db 15 GVISVTVSLLSLGACIGPLMKDAKHWLHFFFLAMAVSLSSDAILHLIIPQVGVGDH 74
QY 383 ASHH-HSHSHPEPAMWKRGPLFSLSSQNIIESAYFDSTW-----KGLTALG----G 430
Db 75 GHNHGHSHHHHEHTDLENG-TSGHGHSHHDDSDRI--MWTVTPARKNLLRLSVLYLS 131
QY 431 LPEMLVEHLTLIKQKDKKKKQKPNDDDDVBIKKQLSKYESQLSTNEEKVDTDR 490
Db 132 IYILPFVEFMFYRKTHLHYCSPSTKSPVITSVSTHSDIKSSASNSHITGSEEDNNDS 191
QY 491 EGYLRADSOEPLSHFDSQOPAVLEEEVYMAHAHPQEVYNEYVPRGCKNCKSHFDTLQ 550
Db 192 SKERRNSVE-----LEKRRQGGH----- 211
QY 551 SDDLHHHDYHHLHHHHQNHHPHSHSQRSREBKD-----AGVATLAWVMINGDL 605
Db 212 --ELIS-----LREDGDDGTETCGKPRALILFGDGV 243
QY 606 HNFSDGLATGAFTGLSSGLSTVAVFCHLPHLPHGLFVALLKAGMTVKQAVLYNLSA 665
Db 244 HNVLDGLANGSFMISVKLFTITIAVICHELPHIGSLAVLIDSGLSMCTAILMLISA 303
QY 666 MLAYLGMARGIFIGHYANVSMWIFALTAGLFMYVALVDM 705
Db 304 LTAYAGLFAIVLGR-DEEIEITAILAITAGMELYVAWDM 342

RESULT 5
T24835
hypothetical protein T11F9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A:Accession: T24835
R:Riennard, N.
submitted to the EMBL Data Library, June 1996

A:Reference number: Z19941
A:Accession: T24835
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-586 <WIL>
A:Cross-references: EMBL:Z74042; PIDN:CAA98527.1; GSPDB:GN00023; CESP:T11F9.2
A:Experimental source: clone T11F9
C:Genetics:
A:Gene: CESP:T11F9.2
A:Map position: 5
A:Introns: 86/3; 123/1; 219/2; 316/3; 538/3

Query Match 9.5%; Score 381.5; DB 2; Length 586;
Best Local Similarity 21.3%; Pred. No. 6.1e-20;
Matches 120; Conservative 110; Mismatches 182; Indels 151; Gaps 13;
QY 286 ERFNYLCPALIIQIDARSCLHTSEKAEI-----PPKTY 319
Db 63 KPFYMLAALNFSISKSIEIYTIKNASQLIHQVLHDAFLADLGNYTEAVNLIDKPPAWQ 122
QY 320 SIOIAWVGFTAIISLISLGLVILPLMNRVFKFLSLFVALAVGLSGDAFLHLLP 379
Db 123 T-----WGIGFAIVSGCSFSAPLGLILLPLCLSKSLYRTIMFLVAVGIALSGSTIFIMP 178
QY 380 HSHASHHHSHSHEEPAMWKRGPLFSLSSQNIIESAYFDSTWGLTALGGLYFVFLVH 439
Db 179 QAF-----HLTS-----FEHFYHKSLLILCALYAFYVDR 210
QY 440 VLTLLIKQFDRKKKKQ-----KKPDDDDVEIKKQLSKYESQLSTNEEK 483
Db 211 MLQYILEFRRRRQTKRRIHASTIASLMTPTTKRRDNGH-----TTEETI 256
QY 484 VTDDETEGILRADSQPSH-----PDSQOPAVLEEEVYMAHAHPQEVYNEYVPRG 535
Db 257 VPSEPTTILQVPLDNNHFRFSELSDEVERTEQKEKMAELANDLEMTNNVYLR 316
QY 536 CKNK-----CHSHFDTLGOSDOLIHSH-----HDYHHILHHHHQ----- 571
Db 317 FSTRRVAVVSGGLDDIEFRSPKLSHHTNGNTSFLOVINNEFRHMTPLSRPQSPVT 376
QY 572 -----NHPHSHSQRSY-----REELKDAG--VATLAWVMINGD 603
Db 377 INVEEPKESYEMKQKSEKPGLNHDNSMSYSIRWEKVIETEPAAVEASVAVMIIFS 436
QY 604 GLHNFSDGLATGAFTGLSSGLSTVAVFCHLPHLPHGLFVALLKAGMTVKQAVLYN 663
Db 437 SANFVDMGSMGAAPSDNLLRGLSLGIIVISQOQFELGLTALILVKSGLGKLTLLFNMV 496
QY 664 SAMLAYLGMARGIFIGHYANVSMWIFALTAGLFMYVALVDMVPEMLHND---ASDHGS 720
Db 497 PIVLSFLGSLGVMLDSVDDSDYDEYIFAISSGMYWIFLGLTILPEIRESTNELIKENLAE 556
QY 721 RWGYFFLQAGMLLIGFGIMLLIS 743
Db 557 SILVSIQACILFGITTFMYFMS 579

RESULT 6
T25420
hypothetical protein T28F3.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
A:Accession: T25420
R:Mortimore, B.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20032
A:Accession: T25420
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-404 <WIL>
A:Cross-references: EMBL:Z82285; PIDN:CAB05297.1; GSPDB:GN00022; CESP:T28F3.3
A:Experimental source: clone T28F3
C:Genetics:

Db 162 -----GSIHS----- 166

QY 569 HQNHHPHSQRYREELKADGAVATLAWMTMGDLNPSDGLAIGAAFTBGLSSGLST 628

Db 167 -----HSHHTPQQTAB--KKAGFNNSAVLNVISGIAHHTDGLATATFYSTQVGIMT 219

QY 629 SVAVFCHPHELGHDFALVAKAGMTKQAVLYNALSAMLAVLGMATGIGTGHY----- 681

Db 220 STAVTFEIPHELGHDFAILSSGFTPPQAIRAQAVTAFGAVGTFSGCWNIGNNSHKA 279

QY 682 ---ANVSMWTFALTAGLNFYALVDMVPEMLHNDASDHGCSRWGYFFLQAGMLLGRGI 738

Db 280 TSSANASBELMPTTAGLLIYIATTSVVPQILHSSAPDSKLRFKKWALQLVFIQVGFAY 339

QY 739 MLLISIEFH 747

Db 340 MALMD--EH 346

RESULT 9

T39240

hypothetical protein SPAP8A3.03 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T39240

R:Wedler, H.; Duesterboeft, A.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, September 1999

A:Reference number: Z21742

A:Accession: T39240

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-453 <WEB>

A:Cross-references: EMBL:AL117210; NID:e1549906; PIDN:CA85170.1; GSPDB:GN00066; SPDB:SF

A:Experimental source: strain 972h; clone p1 p8A3

C:Genetics:

A:Gene: SPDB:SPAP8A3.03

A:Map position: 1

Query Match 6.7%; Score 271.5; DB 2; Length 453;

Best Local Similarity 19.9%; Pred. No. 4.8e-12;

Matches 116; Conservative 94; Mismatches 191; Indels 183; Gaps 17;

QY 179 KDSVASSEVTSTVNTVSEGTHTETPRPKLPKDYSSSTPPSVTSKRVSLRAGR 238

Db 33 RDSFLSQENMKINHET-----TIE-----RLF-REMTENDPILLSKTLAELSGK 77

QY 239 KTVESVSPRGFMYSRNT-----NENPQECFNASKLTHSGMGIQVPLNATEFNYLCPAI 294

Db 78 ELAKAREDLKSVLFLKNNLPVDTESSEAFTEK----- 112

QY 295 INQIDARSLHTSEKRAEIPPKTYSLOIAWGGFIALSIISFLSLGVLVPLMNRVFF 354

Db 113 ----DNNSCVLNSKFEVKQSYSS--SGTNGLIATFATPPNFIILVLP---KSTD 162

QY 355 KFLLSFLVALAVGTLSGDAFLHLLPHSHASHHSHSHEEPAMKRGKPLFSLHSQNIIE 414

Db 163 TSMNLNLFVANSAGSLGDVFLQLPPTVYSTN-----GGDFPASSVYSI- 205

QY 415 SAYFDSWKGTALGGYFMFLVEHVLTLLIKQFDKKKNQKRPENDDDVEIKQLSXYE 474

Db 206 -----LIGALVFFLMDKGRILIH-----RPSLSKPKKDG-----E 238

QY 475 SQLSTNE--EKVTDDBR-TEGYLRADSOEPSHDSQPAVLEE--EYVMAHAHPQEVY 528

Db 239 EYSSVKNKFSASSTQDVKGVEGLRKNRVKNDQNSKGHEFDLIRHVVEV-----SEY 291

QY 529 NEVVPGRCKNKHSHFHDLTGQSDLLIHHDHYHLLHHHHQNHHPHSQRSREELK 588

Db 292 ND----- 293

QY 589 DAGVATLAWMTMGDLNPSDGLAIGAAFTBGLSSGLSTSVAVFCHPHELGHDFAYLL 648

Db 294 ----KTVVTLNLCDSFENMDGLATSAFTNTSIGISTTFAVLLHEIPABIGDLAILL 349

QY 649 KAGMTVKQAVLYNALSAMLAVLGMATGIFI-----CHYAENVSMWIFALTAGLPM 698

Db 350 RNYTKRSQVLYQMTITWYTGLLGAIYATYITASSSSPYGSLQLQLEDLLFFTAGGFL 409

QY 699 YVALVDMVPEMLHNDASDHGCSRWGYFFLQAGMLLGFIMLLI 742

Db 410 XTALGVFPPELLEINLSKGLNWIYALYMMFIVGGSFYLYV 453

RESULT 10

C96704

unknown protein, 23065-20358 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C96704

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alc

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marzi

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venier, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; NUID:21016719; PMID:11130712

A:Accession: C96704

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-450 <STO>

A:Cross-references: GB:AE005173; NID:g6553886; PIDN:AAF16552.1; GSPDB:GN00141

C:Genetics:

A:Gene: T23K23.5

A:Map position: 1

Query Match 6.2%; Score 250.5; DB 2; Length 450;

Best Local Similarity 18.2%; Pred. No. 1.6e-10;

Matches 124; Conservative 91; Mismatches 164; Indels 303; Gaps 22;

QY 101 DBEHSDDHRSDDHSHSDHSDHSDHSHHNAASGNKRRKALCPDHDSDSSGKDPN 160

Db 36 DHVHH--HGGGCGSHS--DHDHDDHDDH----- 60

QY 161 SQCKGAHREHASGRNVKDSVASSEVTSTVNTVSEGTHTETPRPKLPKPDWSS 220

Db 61 -----VKKTTAKVEM----- 70

QY 221 STPPSVTSKRSVRLAGKNTNVSVPKGFMTSRNTNENPQECFNASKLTHSGMGIOV 280

Db 71 -----KLPEELAE-----BEDMRLC-----GFG--- 88

QY 281 PLNATEFNYLCPAIQIDARSLHTSEKRAEIPPKTYSLOIAWGGFIALSIISFLSL 340

Db 89 -----PCL-HDHDHSSSTLTGFAI---WLNALGCSLLVLSL 123

QY 341 LGVILVPLMNR---VFFKFLLSFLVALAVGTLSGDAFLHLLPHSHASHHSHSHEEPAME 397

Db 124 IGVLLPIMPALLIVLEKMLSFVYSCAI-----MQELCWE 159

QY 398 MKRGPLFSLSSQNIIEESAVFDSTWGLTALGGLYFMFLVEHVLTLLI----- 444

Db 160 ML---FFTCNCPMLLVATLITMTMTMTIMILIRLHLLHLYKLCCLDCLFLGLWS 216

QY 445 -----KQFKKKKKKKPENNDD--VEIKQLSKYESQLSTNEEKVTDDBR 492

Db 217 SFLWRSWCSSKKIKDEGHDHNLQSSSDAIVNSSEKYSGGSTDKSLRKKTKASDATD- 275

QY 493 YLRADS--QEPSHDSQPAVLEEEYVMAHAHPQEVYNEVVPGRCKNKHSHFHDLTGQ 550

Db 276 --KDSGTETTSDDKSKDKPEQVET-----R 298

QY 551 SDDLIIHHHDYHLLHHHHQNHHPHSQRSREELKADGAVATLAWMTMGDLNPSD 610

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 21:39:56 ; Search time 25 seconds

(without alignments)

1420.206 Million cell updates/sec

Title: US-09-642-034-5

Perfect score: 4024

Sequence: 1 MARKLSVILLIFALSTVNP.....FGIMLLISFEHKIVIRNF 755

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	455	11.3	449	1 CSUP_DROME	Q9v3a4 drosophila
2	393.5	9.8	515	1 KE4L_CAEEL	Q9xtq7 caenorhabdi
3	389	9.7	476	1 KE4_MOUSE	Q31125 mus musculu
4	376	9.3	404	1 YGUK_CAEEL	Q9xuc4 caenorhabdi
5	355.5	8.8	469	1 KE4_HUMAN	Q92504 homo sapien
6	299	7.4	346	1 YIC3_YEAST	P40544 saccharomyc
7	282.5	7.0	352	1 KE4_BRARE	Q9pnb8 brachydanio
8	273	6.8	355	1 Y816_DROME	Q9vaf0 drosophila
9	168	4.2	351	1 HRPX_PLALO	P04929 plasmodium
10	165.5	4.1	338	1 IARI_ARATH	Q9m647 arabidopsis
11	161.5	4.0	972	1 TOPI_DROME	P30189 drosophila
12	157	3.9	302	1 HYPE_BRAJA	Q45257 bradyrhizob
13	154	3.8	337	1 ZNUA_HAEIN	P44526 haemophilus
14	147.5	3.7	291	1 ZUPT_CAMEJ	Q9pin2 campylobact
15	146	3.6	549	1 DSX_DROME	P23023 drosophila
16	145	3.6	254	1 GUPA_MYXXA	Q06916 myxococcus
17	142	3.5	1709	1 CHDI_HUMAN	O14646 homo sapien
18	141	3.5	950	1 URSL_GSTWA	P40349 ustilago ma
19	140	3.5	332	1 HRPI_PLAFA	P05227 plasmodium
20	140	3.5	732	1 YLH3_SCHPO	Q9hg33 schizosacch
21	139.5	3.5	2026	1 CYAA_YEAST	P08678 saccharomyc
22	137.5	3.4	1711	1 CHDI_MOUSE	P40201 mus musculu
23	135.5	3.4	149	1 EGG3_SCHMA	P08016 schistosoma
24	135	3.4	722	1 HMN2_DROME	P22808 drosophila
25	134	3.3	503	1 YKR5_YEAST	P34240 saccharomyc
26	134	3.3	1093	1 PI4K_DICDI	P54677 dictyosteli
27	133	3.3	657	1 KNOB_PLAFA	P06719 plasmodium
28	131.5	3.3	1402	1 SALM_DROVI	P39806 drosophila
29	130.5	3.2	257	1 ZUPT_ECOLI	P24198 escherichia
30	129.5	3.2	625	1 DALY_DROME	Q24114 drosophila
31	129	3.2	257	1 ZUPT_SALTY	Q8x3r4 salmonella
32	129	3.2	285	1 ZUPT_CLOPE	Q8xmg8 clostridium
33	127.5	3.2	390	1 PC_DROME	P26017 drosophila

RESULT 1					
CSUP_DROME					
ID	CSUP_DROME	STANDARD;	PRT;	449 AA.	
AC	Q9V3A4;				
DT	16-OCT-2001 (Rel. 40, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Catecholamines up protein.				
GN	CATSUP OR CG10449.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephyrdoidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=99403013; PubMed=10471719;				
RA	Sathakis D.G., Burton D.Y., McIvor W.E., Krishnakumar S.,				
RA	Wright T.R., O'Donnell J.M.;				
RT	"The catecholamines up (Catsup) protein of Drosophila melanogaster				
RT	functions as a negative regulator of tyrosine hydroxylase activity.";				
RL	Genetics 153:361-382(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Berkeley;				
RC	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,				
RA	Fosler K.C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,				
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,				
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,				
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,				
RA	Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,				
RA	Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,				
RA	Palazolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,				
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,				
RA	Shue B.C., Siden-Kimms I., Simpson M., Skupski M.P., Smith T.,				
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,				

O15535 homo sapien
O13849 schizosacch
P76425 escherichia
Q08484 saccharomyc
P47395 mycoplasma
Q05738 mus musculu
P19212 neurospora
P28155 rhizobium 1
P13817 plasmodium
P09346 plasmodium
P04930 plasmodium
P21750 drosophila

ALIGNMENTS


```

QY 516 EYMTAHAPQEVNEYVPRGCKNCKCHSHFDITLQSDLLHHHHHHYHHLHHHHQNHHP 575
Db 304 -----
QY 576 HSHSQYSREELKDAGVATLAWMVGDLNFDGLNSDGLAIGAAFTGSLSSGLSTSVAVFCH 635
Db 304 -----LAADFTHFTDGLAIGASFLVGPVAVGAVTTITILLH 339
QY 636 ELPHEIGDFAVLL 648
Db 340 EWPHEIGDFAVLLV 352

RESULT 8
ID Y816 DROME STANDARD; PRT; 355 AA.
AC QSVAF0:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein CG7816.
GN CG7816.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin G.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegvam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE K44/CATSUP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AE003771; AANI4195.1; -.
DR FlyBase; FBgn0039714; CG7816.
DR InterPro; IPR003689; zn_transp_4ip.
DR Pfam; PF02535; zip; 1.
KW Hypothetical protein; Transmembrane; Glycoprotein.
FT TRANSMEM 37 57 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 118 138 POTENTIAL.
FT TRANSMEM 273 293 POTENTIAL.
FT TRANSMEM 301 321 POTENTIAL.
FT CARBOHYD 4 4 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 218 218 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 355 AA; 38870 MW; 01527C0390741FE8 CRC64;
SQ SEQUENCE 355 AA; 6.8%; Score 273; DB 1; Length 355;
Query Match Best Local Similarity 20.8%; Pred. No. 6.8e-12;
Matches 96; Conservative 76; Mismatches 162; Indels 128; Gaps 14;
QY 299 DARSCLHTSEKKAETPPKTYISLQIA-WVGGFATISIIISFLSLGLVILYPL---MNRVVF 354
Db 9 DEHTAMTYSNLMDOQYMFYKFSFEYTPWVFSLLGSVIGLSGIFLLIIPTEEKMAKEG 68
QY 355 K-----FLLSFLVALAVGTLSDAFHLHPHSHASHHSHSEEPAMEMKRGPLFSHLS 409
Db 69 KDPADSKLLRVLLSPFVAGGLGDLVFLHLLPAWEG-----DNQDPS-----SHPSL 114
QY 410 QNIEESAYFSTWKGTLGGLYFMFLYVEHVLILIKOFKDKKKKKPKPENDDDVEIKKQ 469
Db 115 RS-----GLWLSGLIIFTIVEKIFS----- 135
QY 470 LSKYESQLSTNEKVKVTDRTDTEGYLRDSQEPFSDSQPAVLEEEYVIAHAHPQVYN 529
Db 136 -----GYASADENP-----QPKVEIANCEIL-FRGGGLPE 166
QY 530 EYVPRGCKNCKCHSHFDITLQSDLLHHHHHHYHHLHHHHQNHHPHSHSQYSREELKD 589
Db 167 GETSESCGGACDI---EDVGKVCLE-----REQEOKSKERKE 200
QY 590 AGVATLAWMVGDLNFDGLAIGAAFTGSLSSGLSTSVAVFCHPELHSGDFAVLLK 649
Db 201 QPKKVAGYLLNLLANSIDNFTHGLAVAGSFLVSPFRHGLATFATFALLHETPHVEVGDFAILLR 260
QY 650 AGMT---VKQAVLYNALSAMLAYLGMATGIFIGHYAEVNSMMIFALTAGLFWYALVDMV 706
Db 261 SGFSNDAAARQILLTAGAGLLGALVAIGSGVTSMEARTSWLMPFTTAGGFLHALVTVL 320
QY 707 PMLHNDASDHGCSRWGMYFFLQNAQMLLGFGIMLLIS-IPFH 747
Db 321 PDLLEKEERKESTK-----QLLALVFGIALMAVMTLFEH 355

RESULT 9
ID HRPX_PLALO STANDARD; PRT; 351 AA.
AC P04929;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE Histidine-rich glycoprotein precursor.
OS Plasmodium lophurae.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5853;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85061618; PubMed=6095114;
RA Ravetch J.V., Feder R., Pavlovic A., Blobel G.;

```



```

DR EMBL: AC012563; AAG52008.1; ALT_SEQ.
DR EMBL: AY072108; AA159930.1; -.
DR InterPro; IPR003689; zn_transp_tzip.
DR Pfam; PF02535; zip; 1.
KW Transport; Transmembrane; Glycoprotein.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 114 134 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT DOMAIN 37 60 HIS-RICH.
FT DOMAIN 161 242 HIS-RICH.
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 338 AA; 38954 MW; 07095C8B5BC01DAB CRC64;

Query Match
Best Local Similarity 18.7%; Pred. No. 0.00022;
Matches 85; Conservative 49; Mismatches 104; Indels 217; Gaps 20;

QY 101 DHEHSDHERHSDHEHSDHEHSDHSHHNAAGSKKAKALCPDHSDSSGKDPFN 160
Db 36 DHVH--HGCGSHS--DHDHDDHDDH----- 60
QY 161 SQGGAHPHGHASGRNRNVDVSASEVTSVNTVSEGTHTLETPTRECKLPKDVSS 220
Db 61 -----VKTKAKVEM----- 70
QY 221 STPPSVTSKSVRLAGRTNSVSEPRKGFMYSRNINENPQCFNASKLTSHGMGIQV 280
Db 71 -----KLPELAE-----EEDMLC-----GFG--- 88
QY 281 PLNATEFNYLCPALINQIDARSLHTSEKKAETPPKTYISLQTAWVGFGFALSIISPLSL 340
Db 89 -----PCL-HDDHDESSSTLGFAL---WLNAGCSLIVSLASL 123
QY 341 LGVILVPLMNRVFF-----KFLSLFVALAVGTLSGDAFLHLPHS-----HAS 384
Db 124 ICLVLLPIM--FVQCKPSKWFVDSIALFGAGMLGDAFLHQLPFAFGGSHSHNDHEN 180
QY 385 H-HHSHSEEPAMKRGPLFHLSSQNIIESAVFTDWTGLPALGGLYFMFLVHVLTL 443
Db 181 HDHDDHSD-----SPSHS-SIQDLS-----VGLSVLAGIVVFLVLEKLVRY 223
QY 444 I-----KQFKDKKKKKKKPPENDDD--VEIKKOLSKYSQSLSTN 480
Db 224 VEENSSGNTWTHHHHHHAGSKKLKDEGDNHNDQSSSDALVNSSEKVSQGGSDXSRL 283
QY 481 EEKYDTRTEGYLRADS--QPSHFDSQQPAVLE 513
Db 284 KRKTSASDATD--KSDSGTEITSDBGSKDKPEQVE 315

RESULT 11
ID TOPI_DROME STANDARD; PRT; 972 AA.
AC P30189; Q9YXW6;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2).
GN TOPI OR CG6146.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93117086; PubMed=1335568;
RA Hsieh T.-S., Brown S.D., Huang P., Postel J.;
RT "Isolation and characterization of a gene encoding DNA topoisomerase
  I in Drosophila melanogaster.";
  -----
  RL Nucleic Acids Res. 20:6177-6182(1992).
  [2]
  RN SEQUENCE FROM N.A.
  RP STRAIN=Oregon-R;
  RX MEDLINE=96354910; PubMed=8769417;
  RA Zhang C.X., Lee M.P., Chen A.D., Brown S.D., Hsieh T.-S.;
  RT "Isolation and characterization of a Drosophila gene essential for
  early embryonic development and formation of cortical cleavage
  furrows.";
  [3]
  RL J. Cell Biol. 134:923-934(1996).
  [3]
  RN SEQUENCE FROM N.A.
  RP STRAIN=Berkley;
  RX MEDLINE=20196006; PubMed=10731132;
  RA Adams M.D., Gelinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
  RA Anantades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
  RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
  RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
  RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
  RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
  RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
  RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
  RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
  RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
  RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
  RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
  RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
  RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
  RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
  RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
  RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
  RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
  RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
  RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
  RA Klamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
  RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
  RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
  RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
  RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
  RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
  RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
  RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
  RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
  RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
  RA Svirkas R., Tector R., Turner C., Venter E., Wang A.H., Wang X.,
  RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
  RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
  RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
  RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
  RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
  RT "The genome sequence of Drosophila melanogaster.";
  Science 287:2185-2195(2000)
  CC -I- FUNCTION: THE REACTION CATALYZED BY TOPOISOMERASES LEADS TO THE
  CC CONVERSION OF ONE TOPOLOGICAL ISOMER OF DNA TO ANOTHER.
  CC -I- CATALYTIC ACTIVITY: ATP-independent breakage of single-stranded
  CC DNA, followed by passage and rejoining.
  CC -I- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
  CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
  CC RELAX ONLY NEGATIVE SUPERCOILS.
  CC -I- MISCELLANEOUS: WHEN A TOPOISOMERASE TRANSIENTLY BREAKS A DNA
  CC BACKBONE BOND, IT SIMULTANEOUSLY FORMS A PROTEIN-DNA LINK, IN
  CC WHICH A TYROSYL OXYGEN IN THE ENZYME IS JOINED TO A DNA PHOSPHORUS
  CC AT ONE END OF THE ENZYME-SEVERED DNA STRAND.
  CC -I- SIMILARITY: BELONGS TO THE EUKARYOTIC TYPE I TOPOISOMERASE FAMILY.
  -----
  This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/
  or send an email to license@isb-sib.ch).
  -----

```

```
DR EMBL; M74557; AAA28951.1; -.
DR EMBL; U80064; AAC24158.1; -.
DR EMBL; AE003496; AAF48440.1; -.
DR PIR; S35521; S35521.
DR HSP; P11387; I335.
DR FlyBase; FBgn004924; Top1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0009795; P:embryonic morphogenesis; IMP.
DR GO; GO:0002168; P:larval development (sensu Insecta); IMP.
DR GO; GO:0007292; P:ogenesis; IMP.
DR InterPro; IPR001631; Topoisomerase_I.
DR Pfam; PF02919; Topoisomerase_I; 1.
DR Pfam; PF01028; Topoisomerase_I; 1.
DR PRINTS; PR00416; EUTPIGMARASE1.
DR SMART; SM00435; TOPBUC; 1.
DR PROSITE; PS00176; TOPOISOMERASE_I_BUK; 1.
KW Isomerase; Topoisomerase; DNA-binding.
FT DOMAIN 32 39 POLY-HIS.
FT ACT_SITE 930 930 DNA CLEAVAGE (BY SIMILARITY).
FT CONFLICT 40 40 S > H (IN REF. 3).
FT CONFLICT 46 46 S > SSS (IN REF. 3).
FT CONFLICT 201 201 H > Q (IN REF. 3).
SQ SEQUENCE 972 AA; 111688 MW; 3764B8BDEFA30CD CRC64;

Query Match
Best Local Similarity 4.0%; Score 161.5; DB 1; Length 972;
Matches 103; Conservative 79; Mismatches 162; Indels 179; Gaps 22;

QY 68 NLSVE--GFKLLQNTGIDIKRIHHDDHSDHEHSDHEHSHSDEHHS--DHEHHS 124
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 9 NSHIQNGSGEVVQSGV--TTNGHGH--HHHSSSSSSSKKSKKSKDKHRRERHKS 65
QY 125 DHDHSHHSHASCKNKALCPDHDSDSG-----KPRNSQKGARPEHAGSRNVK 179
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 66 SNSSSSSEHKSRRDKR-----HKSSSSSKHKDKDKDGSNSHRGSSSSHKD-K 119
QY 180 DVSASEVTSTVYNTYSEGTHFTETTPRGKLPFKDVSSST--PSPVTSKSRVSLAG 237
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 120 DGSSSSSKHS-----SSGHKSKSKDKERDKDGRGSSSSSRKSSSSSRDKERSSS 173
QY 238 RTNRSVSEPKPGFMYGRNENPNQBCFNASKLLTSHGMGTQVPLNATFNVLCPALINQ 297
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 174 HKSSSSSSSKSKSSSRHSSSS-----SSSSSSSSSSSSSSSSSSSSSSSSSSSS 196
QY 298 IDARSLHTSEKKAETPPKTYSLQIAWVGFIATISLTF--LSLLGVILVPLMNVFFKF 356
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 197 ----SKDRPSVDGVFKPEPVSCQLMHSG-----SVDAFQMQLG----- 233
QY 357 LLSFLVALAVGTLSGDAFLHLLPFSHASHSHSHSEEPAMEMKRGPLFSLSSQNTESA 416
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 234 --SYEAAAAGTNFNG--NAGANYKNGVEESIVDIKK----EESFNLSQAS 280
QY 417 YFDSWKGLTALGGLYF-----FLVEHVLFLIQ----- 446
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 281 SCD-----YSMSQFRADEPPFYVKEQSTAEEDSTNMNPDHDDADEMDDDE 328
QY 447 ----FKDKKKKQKKP-----ENDDDVEIK-----KOLSKYESOLSTNEEKVD 485
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 329 DVLPMARKKKEATRDGMDGMDDDDDDDIPLARKVKKVKIKKESKSKKRVKEEPPSD 388
QY 486 TDRTEGYLRADSQEPSPHDSQPAV-----LEESEV 517
D 11: : : : : : : : : : : : : : : : : : : : : : : : : : : : :
D 389 DYGNVKKPKKKMKKEP-----EPAYSPGKKQKAKAVEEVEV 425

RESULT 12
ID HYBP_BRAJA
AC O45257;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
```

DR	EMBL; U32698; AAC21794.1; -.				
DR	PIR; D64049; D64049.				
DR	TIGR; H10119; -.				
DR	InterPro; IPR006127; SBP_bac_9.				
DR	Pfam; PF01297; SBP_bac_9; 1.				
KW	Transport; Periplasmic; Zinc transport; Zinc; Metal-binding; Signal;				
KW	Complete proteome.				
FT	SIGNAL	1	23		
FT	CHAIN	24	337		
FT				HIGH-AFFINITY ZINC UPTAKE SYSTEM PROTEIN	
FT				ZNUA.	
FT	DOMAIN	115	163		
FT	METAL	57	57		HIS-RICH.
FT	METAL	171	171		ZINC (BY SIMILARITY).
FT	METAL	235	235		ZINC (BY SIMILARITY).
FT	METAL	337	335		ZINC (BY SIMILARITY).
SQ	SEQUENCE	337 AA;	37659 MW;		3DBB45AB8F06FCB CRC64;

[illegible]

Search completed: September 12, 2003, 22:14:45
Job time : 33 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 21:52:06 ; Search time 88 Seconds
(without alignments)
2213.974 Million cell updates/sec

Title: US-09-642-034-5

Perfect score: 4024

Sequence: 1 MARKLSVILLITFALSVTNP.....FGIMLLISIPHEKIVPRNF 755

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_oraganelle:*
9: sp_plage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3969	98.6	749	4 Q13433	Q13433 homo sapien
2	3542	88.0	765	11 Q8C145	Q8C145 mus musculu
3	2533.5	63.0	505	11 Q8R518	Q8R518 mus musculu
4	2280	56.7	433	4 Q8IXR3	Q8IXR3 homo sapien
5	2057	51.1	382	4 Q96HP5	Q96HP5 homo sapien
6	1337	33.2	835	4 Q5ULF5	Q5ULF5 homo sapien
7	902	22.4	381	11 Q8BX42	Q8BX42 mus musculu
8	787.5	19.6	701	5 Q9VSL7	Q9VSL7 drosophila
9	734	18.2	528	6 Q9GKV2	Q9GKV2 macaca fasc
10	703	17.5	528	11 Q8COL2	Q8COL2 mus musculu
11	657.5	16.3	539	4 Q8N6Y3	Q8N6Y3 homo sapien
12	652.5	16.2	535	11 Q9D856	Q9D856 mus musculu
13	652.5	16.2	535	11 Q9D909	Q9D909 mus musculu
14	640	15.9	489	11 Q8VDL0	Q8VDL0 mus musculu
15	632.5	15.7	531	4 Q15043	Q15043 homo sapien
16	616	15.3	529	4 Q9Y3Z1	Q9Y3Z1 homo sapien

17	596	14.8	160	4 Q8NC35	Q8nc35 homo sapien
18	586.5	14.6	462	11 Q91W10	Q91w10 mus musculu
19	583	14.5	460	4 Q5C0K1	Q5c0k1 homo sapien
20	581.5	14.5	462	11 Q9D5V4	Q9d5v4 mus musculu
21	580.5	14.4	462	11 Q9D4Z6	Q9d4z6 mus musculu
22	579	14.4	481	4 Q96BB3	Q96bb3 homo sapien
23	575.5	14.3	519	5 Q9VUI9	Q9vui9 drosophila
24	573.5	14.3	462	11 Q8BTQ3	Q8btq3 mus musculu
25	573	14.2	393	4 Q96SM9	Q96sm9 homo sapien
26	552	13.7	660	11 Q9DAT9	Q9dat9 mus musculu
27	551	13.7	654	6 Q95KA5	Q95ka5 macaca fasc
28	545.5	13.6	647	4 Q96NN4	Q96nn4 homo sapien
29	543.5	13.5	647	4 Q9H6T8	Q9h6t8 homo sapien
30	404	10.0	626	4 Q9NXC4	Q9nxc4 homo sapien
31	396.5	9.9	230	11 Q8CHL4	Q8chl4 mus musculu
32	391	9.7	360	5 Q9XVR4	Q9xvr4 caenorhabdi
33	390.5	9.7	149	4 Q9NXZ2	Q9nxz2 homo sapien
34	382	9.5	512	5 Q22J95	Q22j95 caenorhabdi
35	374.5	9.3	505	5 Q8I4G0	Q8i4g0 caenorhabdi
36	358.5	8.9	157	5 Q9N354	Q9n354 caenorhabdi
37	355	8.8	469	10 Q8H9F6	Q8h9f6 arabidopsis
38	327.5	8.1	129	4 Q96M13	Q96m13 homo sapien
39	321	8.0	368	4 Q96LF0	Q96lf0 homo sapien
40	299	7.4	268	16 Q8XX89	Q8xx89 raistonia s
41	290	7.2	125	4 Q9NSA4	Q9nsa4 homo sapien
42	285	7.1	348	13 Q8AW42	Q8aw42 brachydanio
43	271.5	6.7	453	3 Q5UT11	Q5ut11 schizosacch
44	246	6.1	384	5 Q8IL42	Q8il42 plasmodium
45	234.5	5.8	388	4 Q96H72	Q96h72 homo sapien

ALIGNMENTS

RESULT 1

Q13433 PRELIMINARY; PRT; 749 AA.
AC Q13433;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Estrogen regulated LIV-1 protein.
GN LIV-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Green C., Morgan H.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41060; AAA96258.2; -
DR InterPro; IPR003689; Zn_transpt_2ip.
DR Pfam; PF02535; Zip; 1.
SQ SEQUENCE 749 AA; 84282 NW; 6B48D4E620733F95 CRC64;

Query Match 98.6%; Score 3969; DB 4; Length 749;
Best Local Similarity 99.2%; Pred. No. 5.8e-307;
Matches 749; Conservative 0; Mismatches 0; Indels 6; Gaps 1;
QY 1 MARKLSVILLITFALSVTNPPLHKKAAAFQTTEKISPNWESGINVDLAISTROYHLQOL 60
Db 1 MARKLSVILLITFALSVTNPPLHKKAAAFQTTEKISPNWESGINVDLAISTROYHLQOL 60
QY 61 FRYGNNLSVEGFRKLLQNIQIDTKRIHIHDDHSDHEHHSDEHHSDEHHSDEH 120
Db 61 FRYGNNLSVEGFRKLLQNIQIDTKRIHIHDDHSDHEHHSDEHHSDEHHSDEH 120
QY 121 EHSDDHSHHNSHAASGNKKKALCPDHDSSGKDPNSQKGAHPPEHASGRNVKD 180
Db 121 EHSDDHSHHNSHAASGNKKKALCPDHDSSGKDPNSQKGAHPPEHASGRNVKD 174
QY 181 SVSASEVTSTVNTVSEGTHTLETIETPRGKLFPPKDVSSSTPPSVTSKRSVSLAGRKT 240

```

Db 175 SVSASEVTSTVNTVSEGTHTETETPRPKLPPKDVSSSTPPSVTSKRSVRLRGKT 234
QY 241 NESYSEPRKGFMYSRNTNENPOCFNASKLITSHGMIQVPLNATEFNYLCPALINOIDA 300
Db 235 NESYSEPRKGFMYSRNTNENPOCFNASKLITSHGMIQVPLNATEFNYLCPALINOIDA 294
QY 301 RSLIHTSEKKAETPKTYSLQIAWGGFIAISIIISPLSLGLVILVPLMNRVFKFLISF 360
Db 295 RSLIHTSEKKAETPKTYSLQIAWGGFIAISIIISPLSLGLVILVPLMNRVFKFLISF 354
QY 361 LVALAVGTLSDAFHLPHSHASHHHSHSHEEPAMKRGPLFSLHSONTERSAYFDS 420
Db 355 LVALAVGTLSDAFHLPHSHASHHHSHSHEEPAMKRGPLFSLHSONTERSAYFDS 414
QY 421 TWKGLTALGGLYFMFLVEHYVTLTIKQFKDKKKKKKKPKENDDVEIKKQSKYSQSLTN 480
Db 415 TWKGLTALGGLYFMFLVEHYVTLTIKQFKDKKKKKKKPKENDDVEIKKQSKYSQSLTN 474
QY 481 EKVTDIDRTGYLRADSPSHFSQOPAVLEEEVNIHAHPQVINYEVPRCKNKC 540
Db 475 EKVTDIDRTGYLRADSPSHFSQOPAVLEEEVNIHAHPQVINYEVPRCKNKC 534
QY 541 HSHFHDTLGQSDDLIHHHDYHHLHHHHQNHHPHSHSQYSREELKDAGVATLANMVI 600
Db 535 HSHFHDTLGQSDDLIHHHDYHHLHHHHQNHHPHSHSQYSREELKDAGVATLANMVI 594
QY 601 MGDGLHNSDGLAAGTAFTGLSSGLSTSVAVFCHLPHELGFAYLLKAGMTVKQAVLY 660
Db 595 MGDGLHNSDGLAAGTAFTGLSSGLSTSVAVFCHLPHELGFAYLLKAGMTVKQAVLY 654
QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALYDVMPEMLHNDASDHCS 720
Db 655 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALYDVMPEMLHNDASDHCS 714
QY 721 RMGYFFLQAGMGLLFGIMLLISIFPEHKIVFRINF 755
Db 715 RMGYFFLQAGMGLLFGIMLLISIFPEHKIVFRINF 749

RESULT 2
Q8C145 ID Q8C145 PRELIMINARY; PRT; 765 AA.
AC Q8C145;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Estrogen regulated LIV-1 protein homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK028976; BAC26223.1; -.
SQ SEQUENCE 765 AA; 86380 MW; C6938B9C371377B CRC64;

Query Match 88.0%; Score 3542; DB 11; Length 765;
Best Local Similarity 86.9%; Pred. No. 5 6e-273;
Matches 666; Conservative 44; Mismatches 44; Indels 12; Gaps 5;

QY 1 MARKLSVILITFALSTNPNHELKA-AAPQPTTEKISPNWESGINVDLAISTROYHLQQ 59
Db 1 MATDLSVILITFALSTNPNHELKA-AAPQPTTEKISPNWESGINVDLAISTROYHLQQ 60
QY 60 LFVRYGENNSLSVEGFRKLQIDKIKRIHH-----DHDHSDHSHSDHSHSD 113

```

```

Db 61 LFVRYGENNSLSVEGFRKLQIDKIKRIHH-----DHDHSDHSHSDHSHSD 120
QY 114 HEHSDHSHSDHSHSHHNAASGNKRRKALCPDHDSDSGKDPNSQGGKAHREPHAS 173
Db 121 HEHSDHSHSDHSHSHSHSHSVAGKNNRKAFCPDLDSDNSGNKDPRTSLGKGRPAEHMK 180
QY 174 GRNVKDSVSASVSTVNTVSEGTHTETETPRPKLPPKDVSSSTPPSVTSKRSV- 232
Db 181 GRNVIKESASSVSTVNTVSEGTHTETETPRPKLPPKDVSSSTPPSVTSKRSV- 239
QY 233 --SRLAKRKTNESVSPPRGCFMYSRNTNENPOCFNASKLITSHGMIQVPLNATEFNYL 290
Db 240 RLRLAKRKTNESVSPPRGCFMYSRNTNENPOCFNASKLITSHGMIQVPLNATEFNYL 299
QY 291 CPATINOIDARSCLHT--SEKKAETPKTYSLQIAWGGFIAISIIISPLSLGLVILVPLM 349
Db 300 CPATINOIDARSCLHT--SEKKAETPKTYSLQIAWGGFIAISIIISPLSLGLVILVPLM 359
QY 350 NRVFKFLSLFVALAVGTLSDAFHLPHSHASHHHSHSHEEPAMKRGPLFSLHSS 409
Db 360 NRVFKFLSLFVALAVGTLSDAFHLPHSHASHHHSHSHEEPAMKRGPLFSLHSA 419
QY 410 QNIESAYFTSTWKGLTALGGLYFMFLVEHYVTLTIKQFKDKKKKKKKPKENDDVEIKKQ 469
Db 420 QNIESAYFTSTWKGLTALGGLYFMFLVEHYVTLTIKQFKDKKKKKKKPKENDDVEIKKQ 479
QY 470 LSKYSQSLSTNEKVDRTGYLRADSPSHFSQOPAVLEEEVNIHAHPQVINYEVPRCKNKC 529
Db 480 LSKYSQSLSTNEKVDRTGYLRADSPSHFSQOPAVLEEEVNIHAHPQVINYEVPRCKNKC 539
QY 530 EYVPRGCKNCKSHFHDTLGQSDDLIHHHDYHHLHHHHQNHHPHSHSQYSREELKD 589
Db 540 EYVPRGCKNCKSHFHDTLGQSDDLIHHHDYHHLHHHHQNHHPHSHSQYSREELKD 599
QY 590 AGVATLANMVI MGDGLHNSDGLAAGTAFTGLSSGLSTSVAVFCHLPHELGFAYLLK 649
Db 600 AGVATLANMVI MGDGLHNSDGLAAGTAFTGLSSGLSTSVAVFCHLPHELGFAYLLK 659
QY 650 AGMTVKQAVLYNALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALYDVMPEM 709
Db 660 AGMTVKQAVLYNALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALYDVMPEM 719
QY 710 LHNDASDHGCRGWYFFLQAGMGLLFGIMLLISIFPEHKIVFRINF 755
Db 720 LHNDASDHGCRGWYFFLQAGMGLLFGIMLLISIFPEHKIVFRINF 765

RESULT 3
Q8R518 ID Q8R518 PRELIMINARY; PRT; 505 AA.
AC Q8R518;
DT 01-JUN-2002 (TREMELrel. 21, Created)
DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Endoplasmic reticulum membrane protein.
GN ERMELIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2188618; PubMed=11891044;
RT "Ermelin, an endoplasmic reticulum transmembrane protein, contains the
RL novel HELP domain conserved in eukaryotes.";
DR EMBL; AB071697; BAB86300.1; -.
DR MGI; MGI:2147279; Ermelin.
DR InterPro; IPR003689; zn_transpt_zip.
DR Pfam; PF02535; Zip; 1.
SQ SEQUENCE 505 AA; 56811 MW; E2678B8BDFE2BF2F CRC64;

```

```
Query Match          63.0%; Score 2533.5; DB 11; Length 505;
Best Local Similarity 94.5%; Pred. No. 5.5e-193;
Matches 477; Conservative 13; Mismatches 14; Indels 1; Gaps 1;

QY 252 MYSRNTNENPQECFNASKLLTSHGMIQVPLNATEFNILCPALINQIDARSLIHT-SEK 310
DB 1 MYSRNTNDNIQECFNTTKLLTSHGMSIQALINATEFNILCPALINQIDARACLIHTASE 60

QY 311 KAEIPPKTYISLOANWGGFIAISIIISFLSLLGVILPLMNRVFFFLISFLVALAVGTLS 370
DB 61 KAEIPPKTYISLOANWGGFIAISIIISFLSLLGVILPLMNRVFFFLISFLVALAVGTLS 120

QY 371 GDALFLLPHSHASHHSHSHEEPAMEMKRGPLFSLHSQNTSEESYFDSWTKGLTALGG 430
DB 121 GDALLHLLPHSHASHHSHSHEEPAMEMKRGPLFSLHSQNTSEESYFDSWTKGLTALGG 180

QY 431 LYFMFLVEHVLTLLIKQFKDKKKKKKKPPENDDDVEIKKQSKYESQSLSTNEEKVDTDR 490
DB 181 LYFMFLVEHVLTLLIKQFKDKKKKKKKPPENDDDVEIKKQSKYESQSLSTNEEKVDTDR 240

QY 491 EGYLRADSPSHFDSQOPAVLEEEVMAHAHPQEVYNEYYPGRCKNKCCHSHHFDTLGQ 550
DB 241 EGYLRADSPSHFDSQOPAVLEEEVMAHAHPQEVYNEYYPGRCKNKCCHSHHFDTLGQ 300

QY 551 SDDLHHHHDDYHLLHHHHHQNHPHSHSORYSREELKDAGVATLAWMVLKGDGLHNFSD 610
DB 301 SDDLHHHHDDYHLLHHHHHQNHPHSHSORYSREELKDAGVATLAWMVLKGDGLHNFSD 360

QY 611 GLATGAATFGLSSGLSTSVAVFCHPELHFGDFAVLLKAGMTVKQAVLYNALSAMLAYL 670
DB 361 GLATGAATFGLSSGLSTSVAVFCHPELHFGDFAVLLKAGMTVKQAVLYNALSAMLAYL 420

QY 671 GWATGIFIGHYAENVSMWIFALTACLEMYVALVDMVPEMLHNDASHGCSRWGFFLQNA 730
DB 421 GWATGIFIGHYAENVSMWIFALTACLEMYVALVDMVPEMLHNDASHGCSRWGFFLQNA 480

QY 731 GMLLGFGIMLLISIFEKIVFRINF 755
DB 481 GMLLGFGIMLLISIFEKIVFRINF 505

RESULT 4
Q8IXR3
ID Q8IXR3 PRELIMINARY; PRT; 433 AA.
AC Q8IXR3
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
OS Similar to LIV-1 protein, estrogen regulated.
OC Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039498; AAH39498.1; -.
SQ SEQUENCE 433 AA; 48605 MW; 72B8B90BD0A1867B CRC64;

Query Match          56.7%; Score 2280; DB 4; Length 433;
Best Local Similarity 100.0%; Pred. No. 6.6e-173;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 276 MGIOVPLNATEFNILCPALINQIDARSLIHTSEKKAIEPPKTYISLIQAWGGFIAISII 335
DB 1 MGIOVPLNATEFNILCPALINQIDARSLIHTSEKKAIEPPKTYISLIQAWGGFIAISII 60

QY 336 SFLSLLGVILVPLMNRVFFKLLSFLVALAVGTLSGDATFLLHLLPHSHASHHSHSHEEPA 395
DB 61 SFLSLLGVILVPLMNRVFFKLLSFLVALAVGTLSGDATFLLHLLPHSHASHHSHSHEEPA 120
```

```
QY 396 MEMKRGPLFSLHSQNTSEESAYFDSWTKGLTALGGLYFMFLVEHVLTLLIKQFKDKKKKNQ 455
DB 121 MEMKRGPLFSLHSQNTSEESAYFDSWTKGLTALGGLYFMFLVEHVLTLLIKQFKDKKKKNQ 180

QY 456 KKPENDDDDVEIKKQSKYESQSLSTNEEKVDTDRTEGYLRADSPSHFDSQOPAVLEEE 515
DB 181 KKPENDDDDVEIKKQSKYESQSLSTNEEKVDTDRTEGYLRADSPSHFDSQOPAVLEEE 240

QY 516 EYMTAHAHPQEVYNEYYPGRCKNKCCHSHFHTLQSDDLLHHHHDDYHLLHHHHHQNHP 575
DB 241 EYMTAHAHPQEVYNEYYPGRCKNKCCHSHFHTLQSDDLLHHHHDDYHLLHHHHHQNHP 300

QY 576 HSHSORYSREELKDAGVATLAWMVLKGDGLHNFSDGLAIGAATFGLSSGLSTSVAVFCH 635
DB 301 HSHSORYSREELKDAGVATLAWMVLKGDGLHNFSDGLAIGAATFGLSSGLSTSVAVFCH 360

QY 636 ELPHELGDFAVLLKAGMTVKQAVLYNALSAMLAYLGNATGIFIGHYAENVSMWIFALTAG 695
DB 361 ELPHELGDFAVLLKAGMTVKQAVLYNALSAMLAYLGNATGIFIGHYAENVSMWIFALTAG 420

QY 696 LFMVYVALVDMV 706
DB 421 LFMVYVALVDMV 431

RESULT 5
Q96HP5
ID Q96HP5 PRELIMINARY; PRT; 382 AA.
AC Q96HP5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
OS Hypothetical protein (fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008317; AAH08317.1; -.
InterPro: IPR003689; Zn_transpt_2ip.
DR Pfam; PF02535; Zip; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 382 AA; 43404 MW; AE4AE1C04401CDBC CRC64;

Query Match          51.1%; Score 2057; DB 4; Length 382;
Best Local Similarity 100.0%; Pred. No. 3.1e-155;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 FLHLLPHSHASHHSHSHEEPAMEMKRGPLFSLHSQNTSEESAYFDSWTKGLTALGGLYF 433
DB 1 FLHLLPHSHASHHSHSHEEPAMEMKRGPLFSLHSQNTSEESAYFDSWTKGLTALGGLYF 60

QY 434 MFLVEHVLTLLIKQFKDKKKKKKKPPENDDDVEIKKQSKYESQSLSTNEEKVDTDRTEGY 493
DB 61 MFLVEHVLTLLIKQFKDKKKKKKKPPENDDDVEIKKQSKYESQSLSTNEEKVDTDRTEGY 120

QY 494 LRADSPSHFDSQOPAVLEEEVMAHAHPQEVYNEYYPGRCKNKCCHSHHFDTLGQSD 553
DB 121 LRADSPSHFDSQOPAVLEEEVMAHAHPQEVYNEYYPGRCKNKCCHSHHFDTLGQSD 180

QY 554 LTHHHDDYHLLHHHHHQNHPHSHSORYSREELKDAGVATLAWMVLKGDGLHNFSDGLA 613
DB 181 LTHHHDDYHLLHHHHHQNHPHSHSORYSREELKDAGVATLAWMVLKGDGLHNFSDGLA 240

QY 614 IGAATFGLSSGLSTSVAVFCHPELHFGDFAVLLKAGMTVKQAVLYNALSAMLAYLGM 673
DB 241 IGAATFGLSSGLSTSVAVFCHPELHFGDFAVLLKAGMTVKQAVLYNALSAMLAYLGM 300
```

```

QY 674 TGIFIGHYAENVSMWIFALTAGLTMVVALVDMVPEMLHNDASDHGSRMGYFFLQAGML 733
DB 301 TGIFIGHYAENVSMWIFALTAGLTMVVALVDMVPEMLHNDASDHGSRMGYFFLQAGML 360
QY 734 LGFGIMLLISFEHKIVIRNF 755
DB 361 LGFGIMLLISFEHKIVIRNF 382

RESULT 6
Q9ULF5 PRELIMINARY; PRT; 835 AA.
AC Q9ULF5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypochemical protein KIAA1265 (Fragment).
GN KIAA1265.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039619; PubMed=10574462;
RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XV.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:337-345(1999).
DR EMBL; AB033091; BAA86579.1; -.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR003689; Znf_transpt_2ip.
DR Pfam; PF02535; 2ip; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
KW Hypothetical protein.
FT NON_VER
SQ SEQUENCE 835 AA; 94559 MW; 851503EFC02F1391 CRC64;

Query Match 33.2%; Score 1337; DB 4; Length 835;
Best Local Similarity 35.5%; Pred. No. 1.8e-97;
Matches 317; Conservative 131; Mismatches 241; Indels 204; Gaps 26;

QY 1 MARKLSVILITLFTALSYNPLHELKAAAPQTEK-----ISPWESSGVNDAIASTR 53
DB 9 MTQKCLICLTFFIHCNCHB-EHDGFEALHQRHGMTELEP---SRFSKQAAENEK 64
QY 54 QYHLOQLPYRGXNNLSVGFRLKQNLQIGDKIRIHHD---HSD----- 101
DB 65 KYVIEKLPERYGNGRLSFFGLEKLLTNLGLGRKRVVEINEDLGHDVSHLDILAVQEG 124
QY 102 ---HEHSDHPRH-----SDHEHSDHHEHSDH 126
DB 125 KHFFSHNHQHSHNLNSENQTVTSVSKRNHKCDPEKTEVSVKSDKDHMDHNLRLH 184
QY 127 DH-----HSHNHAAS-----GNKRKALCP 147
DB 185 HRLKHLHLDHNTHFFNDSTPSEGEPSNEPTNKQEQSDVKLPKGRKKKGRKS 244
QY 148 DHSD--SSGKDPNRSQK-----GAHRPEHASGRNRKDSVASEVTSTVNTVSEGT 200
DB 245 NENSEVTPGPPPHNDQGEQYHNRVHKPDR-----VNPGHSHV 285
QY 201 FLB-----TITPPPKLFPKDVSSSTPPSVTSKRSVRSLAGRKINESVERPK 249
DB 286 LPERNGHPGRGHODLPDNEGEL-----RTRKREAPHVKNNAITSLRK 330
QY 250 GFMYSENINENPQCFNASKLLSHGMIQVPLNATEFNILCPALINQIDARSLIHSE 309
DB 331 DL-----NEDDHHHCLNLTQLKYGHGANSPISTDLFTYLCPALLYQIDSLRCLTEHFDK 386

```

```

QY 310 KKAIEPPPKYSI-----QIAWVGGFIALSIISFLSLGLVILPMNRVFFELLSF 360
DB 387 LLVEDINKDNKLVDPEDRANTGASAWTCGLIISITVLSLSLGLVILPIINQCCFELTTF 446
QY 361 LVAVAGTSLGSDAPLHLLPHSHASHHSHHEEPAMEMKRGPLFSLHSQNTIEESAYFDS 420
DB 447 LVAVAGTMSGDALLHLLPHSGQGHDSHQHAH-----GHGSHGHESNKFLEE---YDA 498
QY 421 TWKGLTALGGYFMYFLVHVLTLIKQDKKKKQK--KPENDDVYEIKKQSKYESQLS 478
DB 499 VLKGVVALGGIYLLFIIEHCIRMFKHVKQQRK-QKWFMKQNTIEESTIGRKLSHKLANT 557
QY 479 TNEEKVD-----TDDRTEGYLRADSQEPSHFDSSQPA-----VLEEEVMIHAHPQEV 527
DB 558 PSDMWLQKPLAGTDDSVSEDRINETELTDLEGQESPPKYNILCEEKIIDHSDGL 617
QY 528 YNEIYVPRGCKNKHSHFDTLGQSDLLIHHDYHHILHHHHHONHHPHSHQSOR---YSR 584
DB 618 HT-----IHEHDLAAAAH---NHGENTKTVLRKHNHWHHKKSHSHSGPCHSG 662
QY 585 EELKAGVATLAWVIMVMDGLHNFSDGLAIGAAPTEGLSSGLSTSVATCHELPHGLGDF 644
DB 663 SLDKETGIANIAWVIMVMDGLHNFSDGLAIGAAPTEGLSSGLSTSVATCHELPHGLGDF 722
QY 645 AVLLKAGMTVKQAVLYNALSMALAYLGMATGIFIGHYAENVSMWIFALTAGLTMVVALVD 704
DB 723 AVLLKAGMTVKQAVLYNLLSAMAYLGMATGIFIGHYAENVSMWIFALTAGLTMVVALVD 782
QY 705 MYPEMLHNDH--SDHGSRMGYFFLQAGMLGFGIMLLISFEHKIVIRNF 755
DB 783 MLPEMLHGDGNEBHGFCVPGQFTLQNLGLLFGFALMVIALYEDKIVFDIQF 835

RESULT 7
Q8BX42 PRELIMINARY; PRT; 381 AA.
AC Q8BX42;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Weakly similar to CDNA FLJ32338 FIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK049099; BAC33542.1; -.
SQ SEQUENCE 381 AA; 42435 MW; 9BDE487CD9765E55 CRC64;

Query Match 22.4%; Score 902; DB 11; Length 381;
Best Local Similarity 46.3%; Pred. No. 2.4e-63;
Matches 189; Conservative 67; Mismatches 104; Indels 48; Gaps 13;

QY 369 LSGDAPLHLLPHSHASHHSHHEEPAMEMKRGPLFSLHSQNTIEESAYFDSFWKGLTAL 428
DB 1 MSGDALLHLLPHSGQGHDSHQHTH-----GHGSHGHESKEFLEE---YDAVLKGLVAL 52
QY 429 GGIYFMYFLVHVLTLIKQDKKKKQK--KPENDDVYEIKKQSKYESQLSSTNEKVD- 485
DB 53 GGIYLLFIIEHCIRMFKHVKQQRK-QKWFMKQSTIEESTIGRKLSHKLNSTPDADWLQL 111
QY 486 -----TDDRTEGYLRADSQEPSHFDSSQPAV-----LEEEVMIHAHPQEVYNYPR 534
DB 112 KPLAGTDDSVSEDRINETELTDLEAQEQSPPKNYLGVEEKIM-DHSHSDGLHT----- 165

```

```

QY 535 GCKNCKSH-PHDTLGSDLLHHRDXYH-ILRHHHONHH---PSSHQRYSRREELKO 589
DQ 166 -----IHEHEVHV-----SHNHEDEKAVLRKHSHQWHEHRAHSHGPGCHSSDLKE 213
QY 590 AGVATLAWVIMGDLNFTSDGLAIGAFTBGLSSGLSTSVAVFCHLPHLGLDFAVLK 649
DQ 214 TGTIANIAWVIMGDIHNFSDGLAIGAFAFSAGLTGTTSTIAVFCHELPHLGLDFAVLK 273
QY 650 AGMTVKQAVLYNALSAMAVLGMATGCTIFGTHYAEVNSWTFALTAGLFTVALVDVPM 709
DQ 274 AGMTVKQAVLYNALSAMAVLGMATGCTAVGQYANNITLWTFATAGLFTVALVDVPM 333
QY 710 LHNDL--SDHSCSWGFFTLQAGMLLGFIMLSTIFKHVIRINF 755
DQ 334 LHGDGDHEHGFCEVGFONLGLLGFETMVLVLYEDKIVDFDQF 381

RESULT 8
Q9VSL7 PRELIMINARY; PRT; 701 AA.
AC Q9VSL7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CG6817 protein.
GN FOI OR CG6817.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley.
EX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.A., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Moadery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Ruskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).

```

```

RN SEQUENCE FROM N.A.
RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin J., Banzon K.Y., Beeson K.A., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoc J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.B., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AEO03555; AAF50401.2; -.
DR FlyBase; FBgn0024236; foi.
DR InterPro; IPR003689; Znf_transpt_Zip.
DR Pfam; PF02535; Zip; 1.
SQ SEQUENCE 701 AA; 76766 MW; 23A45DFD6C2E85F0 CRC64;
Query Match 19.68; Score 787.5; DB 5; Length 701;
Best Local Similarity 26.98; Pred No. 7.3e-54;
Matches 218; Conservative 118; Mismatches 206; Indels 267; Gaps 27;
QY 12 TFALSVTPNPLHELKAAAPQPTTEKISPNWESGINVDLAISTROYHLQQLFYRTGENNSLS 71
DQ 67 TFNYSISPPSRREKRHA---GHEHGPTSESRVP-----QITQVYLEKLM---AQDELMM 114
QY 72 VEGFRKLLQNTGIDKI-----KRTHIHDDHSDHSDHSDHSDHSDHSDHSDHSDHSDH 118
DQ 115 SSGFGLLQQLSLSLASGASGEGTCVPGSRVLVHVQPHDH-----HAH-- 158
QY 119 DHEHSDHSDHSHHNAASKN--KKALC---PDHSDSDSGKDPNRSQCKGAHRPEHAS 173
DQ 159 ---HHEEDHSLQNLNCTLLQNTGTTSNVICPSELPANNTHPLCKEAKNF----- 203
QY 174 GRNNYKDSVSASEVTSTVYNTVSEGHFTLETITETPRPGKLPFKDVSSTPSPVTSKRSVS 233
DQ 204 -----TLSD-----KDLL----- 211
QY 234 RLGRKTNESVSEPRKGFMYSRNTNENPOECFNASKLLTSHGMGIQVPLNATEFNVLCPA 293
DQ 212 -----HLCPI 216
QY 294 IINQIDARS--CL-----IHTSEKKAIPPKYTSLOIAVVGFIATISIFLSILGI 344
DQ 217 LYLKLAQGGCICPAISLSDITTELE-ABEKDIPYVVIYAFISVFACGILGLVGA 275
QY 345 LVPLMNRVFFKLSFLVALAVGTLSGDALFLHLLPHSHASHHSHSHSHSHSHSHSHSHSHSH 404
DQ 276 IIPWMSRYKYYIQLVVALAVGTMTGDLALLHLLHSLA-----GQDEGMIM----- 323

```


[illegible]


```

DE Hypothetical 53.8 kDa protein.
GN BC021530.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.; 2002 to the EMBL/GenBank/DBJ databases.
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021530; AAH21530.1; -.
DR MGD; MGI:2384851; BC021530.
DR InterPro; IPR003689; Zn_transpt_zip.
DR Pfam; PF02535; Zip; 1.
KW Hypothetical protein.
SQ SEQUENCE 489 AA; 53754 MW; CLE4EB9284570FC8 CRC64;

Query Match 15.9%; Score 640; DB 11; Length 489;
Best Local Similarity 29.1%; Pred. No. 2.4e-42;
Matches 166; Conservative 93; Mismatches 158; Indels 154; Gaps 18;

QY 219 SSTTPP-SVTS-----KSRVRLAGKTNESVSEPRKGMKSRMT 257
DB 30 SAGLPPLSATSELDMDRYGKNDLSLTQLAKSLDLHLVGVGRDNYVQPKQEG-----82

QY 258 NENPQCFNASKLITSHGGMGIQVPLNATEFNILCPAINQIDARSCLIHTE--KKAETPP 316
DB 83 PRLNSCFSSGDLFAAHLNLSRSQIGASEFQEFCTILQQLDSQAC--TSENQKSEENE 139

QY 317 KTY-----SQTAWGGFTATISIFSLGLVILVPLMNVFFKFLSLVALAVGTLSD 372
DB 140 QTEEGPSPAIEVWGFSLVSLNLASLLGLVLPCTEKAFPSRVLYFTALSIGTLISN 199

QY 373 AFLHLPHSHASHHSHSHEPAMEMKRGPLFSLSSQNIIESAYFDSWTKGLTALGGLY 432
DB 200 ALFQLLPEAFGN-----PQDNIYS-----KSAVVEGGFY 229

QY 433 FMFLVHVLTILKQPKKKKKKKPNDODDVEIKKQSKYESQLSTNEEKVDTDDRTG 492
DB 230 LFFTEKILKML-----LKQNEHHGHNH-----FTESELPKQDEG 269

QY 493 YLR-----ADQEPESH-----FDSQQPAVLEEEVNIHAHPQEVNEYVPRCKNKNCHS 542
DB 270 VTEKLQGLDLMIPQHCHSELGKAPGT--DEKIVNSMSVQDL-----312

QY 543 HFHDTLQSDDLLHHHHDYHILHHHHQHHPHSHSQRSREELKDAGVATLAWYIMG 602
DB 313 -----QASQSACTYWLKGVRYSDICTLAWMITLS 340

QY 603 DGLHNFSDGLAIGAAFTGLSSGLSTSVAVFCHPELPHGLDFAVLLKAGMTVKQAVLYNA 662
DB 341 DGLHNFIDGLAIGASTVSVFQIGISTSVAILCEEPHGLDFVILLNAGMSIQQALFFNF 400

QY 663 LSAMLAYLGMATGFIG-HYAENVSMWIFALTAGLFMVYVALVDMVPEMLHNDASDHGCSR 721
DB 401 LSACCCYLGLAIGLILAGSHFSAN---WIFALAGGMFLIALADFPPEM--NEVCQEDEN 455

QY 722 WGY---FFLQAGMLLGGIMLISIFEKHI 749
DB 456 DSFLVFFVQLNLGTGFSIMLVLTMSQGI 486

```

RESULT 15

```

Q15043 PRELIMINARY; PRT; 531 AA.
AC Q15043;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein KIAA0062 (Fragment).
GN KIAA0062.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Sazuka T., Tanaka A., Sato S.,
RA Seki N., Kavarabayashi Y., Ishikawa K., Tabata S.
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
DR EMBL; D31887; BAA06685.1; -.
DR InterPro; IPR003689; Zn_transpt_zip.
DR Pfam; PF02535; Zip; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 531 AA; 58417 MW; 5AB18EA399CCE2B CRC64;

Query Match 15.7%; Score 632.5; DB 4; Length 531;
Best Local Similarity 30.3%; Pred. No. 1.1e-41;
Matches 159; Conservative 83; Mismatches 147; Indels 127; Gaps 16;

QY 258 NENPQCFNASKLITSHGGMGIQVPLNATEFNILCPAINQIDARSCLIHTE-----SEKKA 312
DB 124 HRNLSTCFSSGDLFAAHLNLSRSQIGASEFQEFCTILQQLDSQAC--TSENQKSEENEQTE 183

QY 313 EIEPKYTSLOIAWGGFTATISIFSLGLVILVPLMNVFFKFLSLVALAVGTLSD 372
DB 184 EGRPSAVEV---WGYGLLCVTVSLCSLGSAYVPMKTKFYKRLLYFTALAIGTLYSN 240

QY 373 AFLHLPHSHASHHSHSHEPAMEMKRGPLFSLSSQNIIESAYFDSWTKGLTALGGLY 432
DB 241 ALFQLLPEAFG-----FNPLEDYVSKSA-----VVEGGFY 271

QY 433 FMFLVHVLTILKQPKKKKKKKPNDODDVEIKKQSKYESQLSTNEEKVDTDDRTG 492
DB 272 LFFTEKILKIL-----LKQNEHHGH-----SHYASESLPSKK-----DQERG 311

QY 493 YLR-----ADQEPESH-----FDSQQPAVLEEEVNIHAHPQEVNEYVPRCKNKNCHS 542
DB 312 VMEKLQGLDLMIPQHCHSELGKAPMW--DEKIVGSLSVQDL-----354

QY 543 HFHDTLQSDDLLHHHHDYHILHHHHQHHPHSHSQRSREELKDAGVATLAWYIMG 602
DB 355 -----QASQSACTYWLKGVRYSDICTLAWMITLS 382

QY 603 DGLHNFSDGLAIGAAFTGLSSGLSTSVAVFCHPELPHGLDFAVLLKAGMTVKQAVLYNA 662
DB 383 DGLHNFIDGLAIGASTVSVFQIGISTSVAILCEEPHGLDFVILLNAGMSIQQALFFNF 442

QY 663 LSAMLAYLGMATGFIG-HYAENVSMWIFALTAGLFMVYVALVDMVPEMLHNDASDHGCSR 721
DB 443 LSACCCYLGLAIGLILAGSHFSAN---WIFALAGGMFLIALADFPPEM--NEVCQED-ER 496

QY 722 WGY---FFLQAGMLLGGIMLISIFEKHI 749
DB 497 KGSILLPIQLNLGTGFSIMLVLTMSQGI 528

```

Search completed: September 12, 2003, 22:16:28
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 22:14:11 ; Search time 29 Seconds
(without alignments)
1101.541 Million cell updates/sec

Title: US-09-642-034-5
Perfect score: 4024
Sequence: 1 MARKLSVILLITFALSVPNP.....FGIMLLISIFEKIVRINF 755

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/iaa/6C_COMB.pep.*
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2205	54.8	431	1	US-08-311-023-2
2	388.5	9.7	417	4	US-09-599-360B-88
3	166	4.1	150	4	US-09-663-600A-196
4	154	3.8	313	3	US-08-686-528A-3
5	154	3.8	313	3	US-09-456-287-3
6	154	3.8	337	3	US-08-686-528A-2
7	154	3.8	337	3	US-09-456-287-2
8	145.5	3.6	189	1	US-08-152-922A-6
9	143.5	3.6	355	2	US-08-758-621-4
10	143.5	3.6	355	3	US-09-107-858-4
11	140	3.5	309	1	US-08-161-406-2
12	138.5	3.4	167	3	US-09-507-323B-13
13	135	3.4	466	4	US-09-134-001C-4749
14	134	3.3	126	4	US-09-663-600A-102
15	134	3.3	339	4	US-09-328-352-6551
16	132	3.3	363	4	US-09-328-352-4930
17	127.5	3.2	448	3	US-09-461-474-8
18	124.5	3.1	474	3	US-09-461-474-10
19	123	3.1	275	4	US-09-134-001C-5540
20	123	3.1	398	3	US-09-461-474-17
21	121	3.0	89	3	US-09-507-323B-12
22	121	3.0	349	3	US-09-461-474-12
23	121	3.0	649	2	US-08-149-097D-37
24	120.5	3.0	680	4	US-09-452-991A-29223
25	120	3.0	368	1	US-08-211-942-17
26	119.5	3.0	726	3	US-09-426-980-2
27	119.5	3.0	726	3	US-09-476-482-2

28 119.5 3.0 726 4 US-09-517-605-6 Sequence 6, Appli
29 118 2.9 2509 2 US-08-149-097D-35 Sequence 35, Appli
30 117.5 2.9 1182 3 US-09-041-886-21 Sequence 21, Appli
31 116 2.9 312 4 US-09-252-991A-24594 Sequence 24594, A
32 116 2.9 1261 3 US-09-208-742-4 Sequence 4, Appli
33 116 2.9 1261 4 US-09-332-295-2 Sequence 2, Appli
34 116 2.9 1261 4 US-09-709-979-2 Sequence 2, Appli
35 115.5 2.9 345 2 US-08-758-621-14 Sequence 14, Appli
36 115.5 2.9 345 3 US-09-107-858-14 Sequence 14, Appli
37 113.5 2.8 339 2 US-08-758-621-2 Sequence 2, Appli
38 113.5 2.8 339 3 US-09-107-858-2 Sequence 2, Appli
39 113 2.8 707 4 US-09-134-001C-2962 Sequence 2962, Ap
40 112.5 2.8 344 2 US-09-134-001C-3524 Sequence 3524, Ap
41 111.5 2.8 765 2 US-08-663-112-2 Sequence 2, Appli
42 111 2.8 535 2 US-08-007-107-2 Sequence 2, Appli
43 111 2.8 901 4 US-09-134-001C-5389 Sequence 5389, Ap
44 110.5 2.7 912 1 US-07-789-915A-8 Sequence 8, Appli
45 110.5 2.7 912 1 US-08-005-002C-8 Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-311-023-2
; Sequence 2, Application US/08311023
; Patent No. 5693465
; GENERAL INFORMATION:
; APPLICANT: MANNING, David Lockwood
; APPLICANT: NICHOLSON, Robert Ian
; APPLICANT: GEE, Julia Margaret
; APPLICANT: GREEN, Christopher Douglas
; TITLE OF INVENTION: METHODS FOR PREDICTING THE BEHAVIOUR OF
; TITLE OF INVENTION: BREAST TUMOURS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Young & Thompson
; STREET: 745 South 23rd Street
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,023
; FILING DATE: 22-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PATCH, Andrew J.
; REGISTRATION NUMBER: Reg. No. 5693465 32,925
; REFERENCE/DOCKET NUMBER: WCM.56
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/521-2297
; TELEFAX: 703/685-0573
; TELEX: 248425
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-311-023-2

Query Match 54.8%; Score 2205; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3e-203;
Matches 415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 328 GFTATISIFSLGVLIVPLMNRVFFKLLSFLVALAVGTLSDAFLLPHSHASHHH 387
|||||
DB 1 GFTATISIFSLGVLIVPLMNRVFFKLLSFLVALAVGTLSDAFLLPHSHASHHH 60
|||||

Query Match 3.6%; Score 145.5; DB 1; Length 189;

[illegible]

RESULT 11

```

US-08-161-406-2
; Sequence 2, Application US/08161406
; Patent NO. 5476785
; GENERAL INFORMATION:
; APPLICANT: Wellens, Thomas
; APPLICANT: Howard, Russell J.
; TITLE OF INVENTION: Recombinant DNA Clone Containing a
; TITLE OF INVENTION: Genomic Fragment of PfhRP-II Gene from Plasmodium
; TITLE OF INVENTION: Falciparum
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/161,406
; FILING DATE: 06-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1173-448P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-161-406-2

```

```

Query Match      3.5%; Score 140; DB 1; Length 309;
Best Local Similarity 27.4%; Pred. No. 5.2e-05;
Matches 31; Conservative 16; Mismatches 42; Indels 24; Gaps 4;

QY 91 HHHH---DHDHSHDHEHSHDHEHSHDHEH---SDHEHSHDHDHSHH-----N 133
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dδ 93 HAHHAADAHAHAADAHAHAADAHAHAHAADAHAHAHAHAHAHAHAHAHAHAHA 152
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 134 HAHSGNKKKALCPDHDSDSGKDPNRSQCKGAHPPEHAGSRNVKDSVASE 186
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Dδ 153 HAAAYAHHAHAADAHAADA-----HHAAYAHHAHAHAADAHAADAHAATD 198
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12

```

US-09-507-323B-13
; Sequence 13, Application US/09507323B
; Patent No. 6274345
; GENERAL INFORMATION:
; APPLICANT: Choi, Jong Hyun
; APPLICANT: Lee, Sang Yup
; APPLICANT: Xu, Zhaochui
; TITLE OF INVENTION: EXPRESSION VECTORS COMPRISING A GENOME

```

```

; TITLE OF INVENTION: ENCODING OMPC DERIVED FROM E.COLI AS CELL SURFACE ANCHORING
;
; TITLE OF INVENTION: MOTIF
;
; FILE REFERENCE: HYLEE39.001AUS
;
; CURRENT APPLICATION NUMBER: US/09/507,323B
;
; CURRENT FILING DATE: 2000-02-18
;
; PRIOR APPLICATION NUMBER: KR 10-1999-0005773
;
; PRIOR FILING DATE: 1999-02-22
;
; NUMBER OF SEQ ID NOS: 13
;
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 13
;
; LENGTH: 167
;
; TYPE: PRT
;
; ORGANISM: Artificial Sequence
;
; FEATURE:
;
; OTHER INFORMATION: Histidine linker.
;
US-09-507-323B-13

```

Query Match 3.4%; Score 138.5; DB 3; Length 167;
Best Local Similarity 30.0%; Pred. No. 2.7e-05;
Matches 45; Conservative 4; Mismatches 38; Indels 63; Gaps 10

Qy	83	GIDKIKRIIHHDD-----HDHHS-----DHEHHS-----DHERUS-----D	113
Db	28	GLDPSGHHHHHSGLDPSGHHHHHSGLDPSGHHHHHSGLDPSGHHHSGLDPSGHH	87
Qy	114	HEHHS-----DHEHHS-----DHDHHS-----HNNHAASGKKKKALCPD--HDS	151
Db	88	HHHHSGLDPSGHHHHHSGLDPSGHHHHHSGLDPSGHHHHHSG-----LDPSGHHHH	141
Qy	152	DSSGKDPSNGQGAHRPEHAGSGRRNVKDS	181
Db	142	HHSGLD-----SGHHHHHSGLEDILOS	165

RESULT 13

```

US-09-134-001C-4749
; Sequence 4749, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4749
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4749

```

Query Match 3.4%; Score 135; DB 4; Length 466;
Best Local Similarity 23.5%; Pred.No. 0.0003;
Matches 50; Conservative 35; Mismatches 99; Indels 28; Gaps 8

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 12, 2003, 22:18:56 ; Search time 58 Seconds
(without alignments)
1899,377 Million cell updates/sec

Title: US-09-642-034-5

Perfect score: 4024
Sequence: 1 MARKUSVILITFALSVTNP.....FGIMLLSIFBHKIVPRNF 755

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/UCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match	Length	ID	Description
1	3969	98.6	749	15	US-10-177-293-264
2	3906	97.1	752	15	Sequence 264, App
3	1972	49.0	397	10	US-10-176-847-52
4	1337	33.2	831	9	Sequence 52, Appl
5	640	15.9	1232	12	US-09-925-300-1531
6	543.5	13.5	647	10	Sequence 1531, Ap
7	543.5	13.5	647	11	Sequence 86, Appl
8	404	10.0	626	15	Sequence 2396, Ap
9	401.5	10.0	737	12	Sequence 16, Appl
10	395.5	9.8	204	15	US-10-097-340-95
11	234	5.8	354	15	Sequence 16, Appl
12	234	5.8	383	15	Sequence 2394, Ap
13	192	4.8	292	9	Sequence 743, App
14	166	4.1	150	12	Sequence 115, App
15	166	4.1	307	11	US-10-050-704-239
					Sequence 37944, A
					Sequence 196, App
					Sequence 95, Appl

16	166	4.1	307	12	US-10-015-387A-95	Sequence 95, Appl
17	166	4.1	307	12	US-10-006-130A-95	Sequence 95, Appl
18	166	4.1	307	12	US-10-199-672-262	Sequence 262, App
19	166	4.1	307	12	US-10-006-172A-95	Sequence 95, Appl
20	166	4.1	307	12	US-10-187-749-262	Sequence 262, App
21	166	4.1	307	12	US-10-194-457-262	Sequence 262, App
22	166	4.1	307	12	US-10-184-642-262	Sequence 262, App
23	166	4.1	307	12	US-10-196-747-262	Sequence 262, App
24	166	4.1	307	12	US-10-015-392A-95	Sequence 95, Appl
25	166	4.1	307	12	US-10-017-253A-95	Sequence 95, Appl
26	166	4.1	307	12	US-10-173-689-262	Sequence 262, App
27	166	4.1	307	12	US-10-173-690-262	Sequence 262, App
28	166	4.1	307	12	US-10-173-691-262	Sequence 262, App
29	166	4.1	307	12	US-10-173-692-262	Sequence 262, App
30	166	4.1	307	12	US-10-173-694-262	Sequence 262, App
31	166	4.1	307	12	US-10-173-698-262	Sequence 262, App
32	166	4.1	307	12	US-10-173-699-262	Sequence 262, App
33	166	4.1	307	12	US-10-173-707-262	Sequence 262, App
34	166	4.1	307	12	US-10-174-569-262	Sequence 262, App
35	166	4.1	307	12	US-10-174-583-262	Sequence 262, App
36	166	4.1	307	12	US-10-174-587-262	Sequence 262, App
37	166	4.1	307	12	US-10-174-589-262	Sequence 262, App
38	166	4.1	307	12	US-10-174-591-262	Sequence 262, App
39	166	4.1	307	12	US-10-175-736-262	Sequence 262, App
40	166	4.1	307	12	US-10-175-742-262	Sequence 262, App
41	166	4.1	307	12	US-10-175-744-262	Sequence 262, App
42	166	4.1	307	12	US-10-175-745-262	Sequence 262, App
43	166	4.1	307	12	US-10-175-748-262	Sequence 262, App
44	166	4.1	307	12	US-10-175-751-262	Sequence 262, App
45	166	4.1	307	12	US-10-175-754-262	Sequence 262, App

ALIGNMENTS

RESULT 1

US-10-177-293-264
; Sequence 264, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lilie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavaru, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMEN
; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT FILING DATE: 2002-06-21
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05

```
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 749
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-177-293-264

Query Match      98.68; Score 3969; DB 15; Length 749;
Best Local Similarity 99.28; Pred. No. 0;
Matches 749; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

QY 1 MARKLSVILITLTFALSVTNPLHELKAAAPQOTEKISPNWESGINVDLAISTROYHLQOL 60
DB 1 MARKLSVILITLTFALSVTNPLHELKAAAPQOTEKISPNWESGINVDLAISTROYHLQOL 60
QY 61 FRYGENNSLSVEGFRKLLQNIQIDKIKRIHHDHSDHSDHEHSDHEHSDHSDHSDH 120
DB 61 FRYGENNSLSVEGFRKLLQNIQIDKIKRIHHDHSDHSDHEHSDHSDHSDHSDHSDH 120
QY 121 EHHSDHSDHSHNHAASGKNRKALCPDHDSDSGKDPNSQKGARPEHAGSRNRYKD 180
DB 121 EHHSDHSDHSHNHAASGKNRKALCPDHDSDSGKDPNSQKGARPEHAGSRNRYKD 180
QY 181 SVSASEVTSTVNTVSGTHFLETETPRPKLFPKDVSSSTPPSVTSKRSVRLAGRKT 240
DB 175 SVSASEVTSTVNTVSGTHFLETETPRPKLFPKDVSSSTPPSVTSKRSVRLAGRKT 234
QY 241 NESVSEPRKGFMYSRNNENPQECFNASKLLTSHGNGIQVPLNATEFNLCPALINQIDA 300
DB 235 NESVSEPRKGFMYSRNNENPQECFNASKLLTSHGNGIQVPLNATEFNLCPALINQIDA 294
QY 301 RSLCHTSKKAETPPKTYSLQIAWVGGFIAISISFLSLGLVILVPLMNRVFFKLLSF 360
DB 295 RSLCHTSKKAETPPKTYSLQIAWVGGFIAISISFLSLGLVILVPLMNRVFFKLLSF 354
QY 361 LVALAVGTLSGDAFLHLLPHSHASHSHSHSHSHEEPAMEMKRGPLFSLSSONTEESAYFDS 420
DB 355 LVALAVGTLSGDAFLHLLPHSHASHSHSHSHSHEEPAMEMKRGPLFSLSSONTEESAYFDS 414
QY 421 TWKGLTALGGLYFMFLVEHVLTILKQKDKKKKKKKPPENDDDVEIKKQSKYESOLSTN 480
DB 415 TWKGLTALGGLYFMFLVEHVLTILKQKDKKKKKKKPPENDDDVEIKKQSKYESOLSTN 474
QY 481 EEKVDTRTEGILRADSQBPESHFDQQPAVLEEEVMIATAHAPQEVYNEYVPRGCKNKC 540
DB 535 HSHFHTLQCSDDLHGHHDYHLLHHHHQNHHPHSHSQRYSREELKDAGVATLAWMYI 594
QY 601 MGDLNHFSDGLAIGAAFTGELSSGLSTSVAVFCHLPHLGDPAVLLKAGMTVKQAVLY 660
DB 595 MGDLNHFSDGLAIGAAFTGELSSGLSTSVAVFCHLPHLGDPAVLLKAGMTVKQAVLY 654
QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALVDMVPEMLHNDASDHGCS 720
DB 715 RWGYFFLQAGMLLGFGLIMLLISIFEHKIVFRINF 749

RESULT 2
US-10-176-847-52
; Sequence 52, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Velody, Petter Ole
```

```
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAS
; TITLE OF INVENTION: AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 752
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-176-847-52

Query Match      97.18; Score 3906; DB 15; Length 752;
Best Local Similarity 99.28; Pred. No. 0;
Matches 736; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MARKLSVILITLTFALSVTNPLHELKAAAPQOTEKISPNWESGINVDLAISTROYHLQOL 60
DB 1 MARKLSVILITLTFALSVTNPLHELKAAAPQOTEKISPNWESGINVDLAISTROYHLQOL 60
QY 61 FRYGENNSLSVEGFRKLLQNIQIDKIKRIHHDHSDHSDHEHSDHSDHSDHSDHSDH 120
DB 61 FRYGENNSLSVEGFRKLLQNIQIDKIKRIHHDHSDHSDHEHSDHSDHSDHSDHSDH 120
QY 121 EHHSDHSDHSHNHAASGKNRKALCPDHDSDSGKDPNSQKGARPEHAGSRNRYKD 180
DB 121 EHHSDHSDHSHNHAASGKNRKALCPDHDSDSGKDPNSQKGARPEHAGSRNRYKD 174
QY 181 SVSASEVTSTVNTVSGTHFLETETPRPKLFPKDVSSSTPPSVTSKRSVRLAGRKT 240
DB 175 SVSASEVTSTVNTVSGTHFLETETPRPKLFPKDVSSSTPPSVTSKRSVRLAGRKT 234
QY 241 NESVSEPRKGFMYSRNNENPQECFNASKLLTSHGNGIQVPLNATEFNLCPALINQIDA 300
DB 235 NESVSEPRKGFMYSRNNENPQECFNASKLLTSHGNGIQVPLNATEFNLCPALINQIDA 294
QY 301 RSLCHTSKKAETPPKTYSLQIAWVGGFIAISISFLSLGLVILVPLMNRVFFKLLSF 360
DB 295 RSLCHTSKKAETPPKTYSLQIAWVGGFIAISISFLSLGLVILVPLMNRVFFKLLSF 354
QY 361 LVALAVGTLSGDAFLHLLPHSHASHSHSHSHEEPAMEMKRGPLFSLSSONTEESAYFDS 420
DB 355 LVALAVGTLSGDAFLHLLPHSHASHSHSHSHEEPAMEMKRGPLFSLSSONTEESAYFDS 414
QY 421 TWKGLTALGGLYFMFLVEHVLTILKQKDKKKKKKKPPENDDDVEIKKQSKYESOLSTN 480
DB 415 TWKGLTALGGLYFMFLVEHVLTILKQKDKKKKKKKPPENDDDVEIKKQSKYESOLSTN 474
QY 481 EEKVDTRTEGILRADSQBPESHFDQQPAVLEEEVMIATAHAPQEVYNEYVPRGCKNKC 540
DB 475 EEKVDTRTEGILRADSQBPESHFDQQPAVLEEEVMIATAHAPQEVYNEYVPRGCKNKC 534
QY 541 HSHFHTLQCSDDLHGHHDYHLLHHHHQNHHPHSHSQRYSREELKDAGVATLAWMYI 600
DB 535 HSHFHTLQCSDDLHGHHDYHLLHHHHQNHHPHSHSQRYSREELKDAGVATLAWMYI 594
QY 601 MGDLNHFSDGLAIGAAFTGELSSGLSTSVAVFCHLPHLGDPAVLLKAGMTVKQAVLY 660
DB 595 MGDLNHFSDGLAIGAAFTGELSSGLSTSVAVFCHLPHLGDPAVLLKAGMTVKQAVLY 654
QY 661 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALVDMVPEMLHNDASDHGCS 720
DB 655 NALSAMLAYLGMATGIFIGHYAENVSMWIFALTAGLFMYVALVDMVPEMLHNDASDHGCS 714
QY 721 RWGYFFLQAGMLLGFGLIMLLI 742
DB 715 RWGYFFLQAGMLLGFGLIMLLI 736

RESULT 3
US-09-925-300-1531
```

; Sequence 1531, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P2101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1531
; LENGTH: 397
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (179)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (180)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (181)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (358)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1531

Query Match 49.0%; Score 1972; DB 10; Length 397;
Best Local Similarity 98.7%; Pred. No. 1.5e-164;
Matches 376; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 181 SVASAEVTSTVNTVSEGTHTLEIETPRPKLFPKDVSSSTPSVTSKSRVRLAGRT 240
|||||
Db 1 SVASAEVTSTVNTVSEGTHTLEIETPRPKLFPKDVSSSTPSVTSKSRVRLAGRT 60
QY 241 NESVPRKGMYSRNTNENPQECFNASKLLTSHGMGIQVPLNATFEFNYLCPALINQIDA 300
|||||
Db 61 NESVPRKGMYSRNTNENPQECFNASKLLTSHGMGIQVPLNATFEFNYLCPALINQIDA 120
QY 301 RSLIHTSEKKAIEPPKTYSLQIAWVGGFIAISIFLSLLGVILVPLMNRVFFKFLSF 360
|||||
Db 121 RSLIHTSEKKAIEPPKTYSLQIAWVGGFIAISIFLSLLGVILVPLMNRVFFKFLSX 180
QY 361 LVALAVGTLSGDAFLHLLPHSHASHSHSHEEPAMKRGPLFSLHSSQNIESAYFDS 420
Db 181 XVALAVGTLSGDAFLHLLPHSHASHSHSHEEPAMKRGPLFSLHSSQNIESAYFDS 240
QY 421 TWKGLTALGGLYFMFLVEHVLTLLKQFKDKKKKNQKPPNDODVEIKKQSKYESOLSTN 480
Db 241 TWKGLTALGGLYFMFLVEHVLTLLKQFKDKKKKNQKPPNDODVEIKKQSKYESOLSTN 300
QY 481 EEKVYDTRTEGYLRADSQSPSHFDSQQPAVLVEEVEVMTAHAPQEVYVNEVPRGCKNC 540
Db 301 EEKVYDTRTEGYLRADSQSPSHFDSQQPAVLVEEVEVMTAHAPQEVYVNEVPRGCKXC 360
QY 541 HSHPHDTLQSDDLTHHHHDY 561
Db 361 HSHPHDTLQSDDLTHHHHDF 381

RESULT 4

US-09-789-561-86
; Sequence 86, Application US/09789561
; Patent No. US20020064818A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.

; TITLE OF INVENTION: 52 Human secreted proteins
; FILE REFERENCE: P2043F1
; CURRENT APPLICATION NUMBER: US/09/789,561
; CURRENT FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: PCT/US00/24008
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 60/152,317
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/152,315
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 86
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-789-561-86

Query Match 33.2%; Score 1337; DB 9; Length 831;
Best Local Similarity 35.5%; Pred. No. 2.7e-108;
Matches 317; Conservative 131; Mismatches 241; Indels 204; Gaps 26;

QY 1 MARKLSVILLILFALSVTNPLHELKAAAFPTTEK-----ISPWESGINVDLAISRP 53
Db 5 MHTKFLICLLIFIPHHCHNCHE-EHDHGPEALHROHGMTELEP---SKFSQAAENK 60
QY 54 QYHLOQLFYRYGENSLSVGEGRKLLQNIQIDKIRIHHD---HDI--HSD----- 101
Db 61 KYVTEKLFERYGENSLRFFGLEKLLNLGLGRKVVIEINHEDLGHSHLODLAVQBG 120
QY 102 ---HEHSDHERH-----SDEHSHSDHEHSDH 126
Db 121 KHFSHNHSHHNLNSENQTVTSVTKRHKODEKETVSVSKSDKHMDHNRHLR 180
QY 127 DH-----HSHNHAAS-----GKNRKALCP 147
Db 181 HRLHLLHLDHNTTHFHNDSTITPSEGEPSNEPSTETNKTOEQSDVKLPKGRKKGRKS 240
QY 148 DHSD--SSGKOPRNSQK-----GAHRPEHAGRRNWKDSVASEVTSTVYTVSEGVH 200
Db 241 NENSEVITPGFPFNHDQGEQYHRNVHKDPR-----VHNPFGSHVH 281
QY 201 FLE-----TIETPRKGLFPKDVSSSTPSVTSKSRVSLAGRKTNESVSPRK 249
Db 282 LPERNGHDPGRGHQDLDPDNEGEL-----RHTRKEAPHVKNNAILSLRK 326
QY 250 GFMYSRNTNENPQECFNASKLLTSHGMGIQVPLNATFEFNYLCPALINQIDARSCLHTSE 309
Db 327 DL---NEDDHHHECLNVTQLKYHGHGANSPISTDLFTYLCALLYQIDSLRCLIEHFDK 382
QY 310 KKAIEPPKTYSL-----QIAWVGGFIAISIFLSLLGVILVPLMNRVFFKFLSF 360
Db 383 LLVEDINKKNIVPDEANIGASAWICGIIISITVLSLLGVILVPLINGOCFKFLTF 442
QY 361 LVALAVGTLSGDAFLHLLPHSHASHSHSHEEPAMKRGPLFSLHSSQNIESAYFDS 420
Db 443 LVALAVGTMSGDALHLLPHSQGGHDHSHQHAF-----GHGSHSHESNKYLEE---YDA 494
QY 421 TWKGLTALGGLYFMFLVEHVLTLLKQFKDKKKKNQK--KPENDDVEIKKQSKYESQLS 478
Db 495 VLKGLVALGGLYLLFTIEHCIRNFKYKQQRGK-QKWFMKQNTTESTIGRKLSDHKLNT 553
QY 479 TNEKVD-----TDDRTEGYLRADSQSPSHFDSQQPA-----VLEEEVMTAHAPQEV 527
Db 554 PDSWLQLKPLAGTDDSVVSEDLNLETETDLQEQSPPKNYLCIEEEKIIDHSHSDGL 613
QY 528 YNEVPRGCKNKHSHFHDTLGQSDDLTHHHHDYHLLHHHHHONHHHPHSHSQR---YSR 584
Db 614 HT-----THEHDLHAAH---NHHGENKTVLRKHNHONHHHSHSGHSGCHSG 658
QY 585 EEKLDAGVATLAWVIMGDLHNFSDGLAIGAAGFTBGLSGSLSTSVAVFCHPELHGD 644
Db 659 SDLKETGIANIWMVIMGDIHNFSDGLAIGAAGFTBGLSGSLSTSVAVFCHPELHGD 718

```

QY 571 QNHPPHSHSQRYSREELKDAGVATLAWVMIMGDLHNFSDDGLAIGATPTEGLSSGLSTSV 630
Db 344 --HQGHSGHQ-----GGTDITWMLLGDGLHNLTDGLAIGATPDSGFSGLSTTL 392
QY 631 AVECHELPHELGDPAVLLKAGMTVQAVLYNALSAMLAYLNATGTFITGHYAENVSMWIF 690
Db 393 AVFCEHLPHELGDPAFLMLOGSLGSRFRLLLSLVSGALGCGAVLGVLGSLGPEVLTFWVF 452
QY 691 ALTAGLFMYALVDMYPEMLHND---ASDHGCSRWGFFFLQAGMLLGFGLMLLSIFEH 747
Db 453 GVTAGVELYALVMDLPAIRPEPLTPPH-----VLIQGLILGLLGLMLAITLLEE 505
QY 748 KIV 750
Db 506 RLL 508

RESULT 6
US-09-965-529-16
; Sequence 16, Application US/09965529
; Publication No. US20020182671A1
; GENERAL INFORMATION:
; APPLICANT: IAL, Preeti
; APPLICANT: YUE, Henry
; APPLICANT: TANG, Y. Tom
; APPLICANT: BANDMAN, Olga
; APPLICANT: BURNFORD, Neil
; APPLICANT: AZIMZAL, Yaida
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: LU, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PR-0731 USA
; CURRENT APPLICATION NUMBER: US/09/965,529
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/149,641; 60/164,203; PCT/US00/22315
; PRIOR FILING DATE: 1999-08-17; 1999-11-09; 2000-08-14
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20020182671A1 1336728CDI
US-09-965-529-16

```

Query Match	13.5%;	Score 543.5;	DB 10;	Length 647;
Best Local Similarity	26.5%;	Pred. No. 8.9e-39;		
Matches 178;	Conservative 99;	Mismatches 201;	Indels 193;	Gaps 23;
QY	130	SHENRASGKKEKALCPD-----HDSGSGKDPNS-----	QKGARREHAS	173
		: : : : : : :		
Db	121	SHADLLALLESKALTPLGSLWLMORMQARAAGTPTACVDIPOLLEAVAGAGPESG	180	
QY	174	GRNVKDSVASWTSTVNTVNSGTHFTETWTPREGKLPKDVSSSTPPSVTSK--SR	231	
		: : : : : : :		
Db	181	G-----VLAALLDHWGSGCF-HALSPQYFVDFVFOQHSSEVPMTLABLSAL	227	
QY	232	VSRL-AGRKNTESVSPRKG-----PMYSRNTNENPQECFNASKLLTSHGMGIQVP	281	
		: : : : : : :		
Db	228	MORLVGVRGAHSDHSHRHGASRRDPVLISSSNSSSVWDTVCLSDRVMAAYGLSEQA	287	
QY	282	LNATEFNYLCPALINQIDARSLTHTSEKAEPTPKTYSLOIA--WVGGFIALISIFLS	339	
		: : : : : : : :		
Db	288	VTPENAWQSPALLOQQLSGAC---TSQSR---PPVQDLQSQSBRYLYGSLATFLITCLCA	341	
QY	340	LLGVILYPLMN-RVFEKFLISFLVALAVGTLSDGAFLHLHPHSHASHHHSHSEEPAMEN	398	
		: : : : : : :		
Db	342	VFGELLTCTGCGVGVTHLTQTFUSLAVGALTGDVVLHLPTRVIGLHTHS-----	391	
QY	399	KRGFLFSLSSQNIBESAYFDSTWKGTLATGLGLYFMFAVEHVLTLIFQFDKDKKKKKKPP	458	

```

Db      392  -----EGLSPQTRWLLAMLAGIYAFFLFENLNL-----LP 425
QY      459  ENDDVVE-----IKQLSKYESOLSNEEKVDTDRREGYLADSP 501
Db      426  RDPEDLEDCGCHSHSHGGSHGVSILQAPSE-----LR--QPKP 464
QY      502  SHFDSQPAVLEEEVMTAHAPQEVYVYVPRGCKNCHSHPHDTLGGSDLLIHHHDY 561
Db      465  PHGSRADIVAE-----SPELL-----482
QY      562  HHILHHHQNHHPHSHSQRYSRRELKDAGVATLAWVMIMGDLHNFSDGLAIGRAFT 621
Db      483  -----NPEP-----RRLSPE-----LRLLPYMTITLGDVAFNEADGLAVGAPASS 522
QY      622  LSSGLSTVAVFCHHELPHLGDFAVLLKAGMTVKQAVLYNALSAMLAYLGMATGIFIGHY 681
Db      523  WKTGLATSLAVFCHHELPHLGDFAVLLHAGLSVROALLNLASALTAFAGLYVALAVG-V 581
QY      682  AENVSMWIFALTAGLFWYVALVDMVPEMLH-NDASDHGCSRWGYFFLQNAAGMLLGFIML 740
Db      582  SESEAWILLAVATGLFYLVALCDMLFAMLKVRDPRP-----WLLFLHNVLGGTGVILL 636
QY      741  LISIFEKIVF 751
Db      637  LLSLYEDDITF 647

RESULT 7
US-09-969-680A-16
; Sequence 16, Application US/09969680A
; Publication No. US20030124649A1
; GENERAL INFORMATION:
; APPLICANT: LAL, Preeti; YUE, Henry
; APPLICANT: TANG, Y. Tom; BANDMAN, Olga
; APPLICANT: BURFORD, Neil; AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.; LU, Dyoung Aina M.
; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA
; CURRENT APPLICATION NUMBER: US/09/969, 680A
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US00/22315
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/149, 641
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/164, 203
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PERL Program
; SEQ ID NO 16
; LENGTH: 647
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030124649A1 1336728CD1
US-09-969-680A-16

Query Match      13.5%; Score 543.5; DB 11; Length 647;
Best Local Similarity 26.5%; Pred. No. 8.9e-39;
Matches 178; Conservative 99; Mismatches 201; Indels 193; Gaps 23;

QY      130  SHINHAASKNKKALCPD-----HDSGSKGPRNS-----QKGHAPREHAS 173
Db      121  SHADHLLALLESFKALTPGLSNLIQRMQARAAGTQPKACVDIPQLLEEAVGAGAPGS 180
QY      174  GRNVKDSVSAEVTSTVNTVSGTHLETETPRPKLPKDVSSSTPPSVTSK--SR 231
Db      181  G-----VLAALLDHVRGSCF-HAUPSPQYFVDFVFOHSSSEVPMTLBSAL 227
QY      232  VSRL-AGRKTNSVSEPRKG-----FMYSRNTNENPQCFNASKLLTSHGMGIQVP 281

```

```

Db      228  MQLGVGREAHSDHSHRGASSRDPVPLISSSSSSVMDTVCLSDARDVMAATGLSEQAG 287
QY      282  LNATEPNYLCPAINQIDARSLIHTSEKKAELPKTYSLQIA--WVGFIALSIISFLS 339
Db      288  VTEPAWAQLSPALLQOOLSGAC---TSQSR---PPVQDLSQSRYLGSLATLCLCA 341
QY      340  LLGVILVPLMN-RYFFKFLLSFELVALAVGTLSGDAFLHLLPHSHASHSHSHEEPAM 398
Db      342  VFGLLLLTCTGVRVTHIILQTFSLAVGALGTDAVHLHLPKVLGLTHS-----391
QY      399  KRGLPFLSHLSSONIBESAYFDSTWKGITLALGGLYFMFLVEHVLTLIKQFKDKKKKQKP 458
Db      392  -----EGLSPQTRWLLAMLAGIYAFFLFENLNL-----LP 425
QY      459  ENDDVVE-----IKQLSKYESOLSNEEKVDTDRREGYLADSP 501
Db      426  RDPEDLEDCGCHSHSHGGSHGVSILQAPSE-----LR--QPKP 464
QY      502  SHFDSQPAVLEEEVMTAHAPQEVYVYVPRGCKNCHSHPHDTLGGSDLLIHHHDY 561
Db      465  PHGSRADIVAE-----SPELL-----482
QY      562  HHILHHHQNHHPHSHSQRYSRRELKDAGVATLAWVMIMGDLHNFSDGLAIGRAFT 621
Db      483  -----NPEP-----RRLSPE-----LRLLPYMTITLGDVAFNEADGLAVGAPASS 522
QY      622  LSSGLSTVAVFCHHELPHLGDFAVLLKAGMTVKQAVLYNALSAMLAYLGMATGIFIGHY 681
Db      523  WKTGLATSLAVFCHHELPHLGDFAVLLHAGLSVROALLNLASALTAFAGLYVALAVG-V 581
QY      682  AENVSMWIFALTAGLFWYVALVDMVPEMLH-NDASDHGCSRWGYFFLQNAAGMLLGFIML 740
Db      582  SESEAWILLAVATGLFYLVALCDMLFAMLKVRDPRP-----WLLFLHNVLGGTGVILL 636
QY      741  LISIFEKIVF 751
Db      637  LLSLYEDDITF 647

RESULT 8
US-10-097-340-95
; Sequence 95, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: JOHN MONAHAN
; APPLICANT: Manjula GANNAVAPURU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumel ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26

```



```

; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 95
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-95

```

Query Match	10.0%;	Score 404;	DB 15;	Length 626;		
Best Local Similarity	24.7%;	Pred. No. 1.5e-26;				
Matches 146;	Conservative 86;	Mismatches 173;	Indels 186;	Gaps 20;		
QY	130	SHHNAASGNKKRKA	CPD-----HDSGSGKDP	RNS-----QKGGAHPHNAS 173		
DB	96	SHADHLIALLEPKAL	TGCLSWLLQRMQARA	AGQTPKTACVDIPOLLEEA	VAGVAGPGSAG 155	
QY	174	GRANVKDYSASEVT	SYVYNYVSGTHFLE	TETRPCKLFPKDV	SSSTPSTVTSK--SR 231	
DB	156	G-----VLAALLD	HVRGSGCF-HAIP	SPOYFVDFVFOH	SEVPMFLAELSAL 20	
QY	232	VSRL-AGRKTIN	SVSEPRKG-----FMS	SRNTNENPOCF	NASKLITSHGMGIQVP 281	
DB	203	MQRLGVGREAH	SDHSHRUGASSR	DPVPLISSNSSSV	YMDTVCLSDARVMAAYGLSEQAG 262	
QY	282	LNATERNYCPAL	INQIDARSLHTT	SEKKAETPKTY	SLQIA--WVGGFIALSIISPLS 339	
DB	263	VTPEWAQIS	PALLOQ---QLSGAY	TSCSR---PPVQD	LSQSERYLYGSLATLLCICA 316	
QY	340	LLGVILVPLMN-R	VYFKFELLSF	VALAVGFLGDA	FLHLPHSHASHHHSHSEEPAMD 398	
DB	317	VFGLLILCTCG	RGVAHYILOTL	SLSLVAGALT	GDVYLHLPKVLGLTHS-----366	
QY	399	KRGFLSHLSQ	NIEESAYFDS	TKWGLTALGG	LYFMFLVEHVLTLLIQFDK	KKKKKKPK 458
DB	367	-----EGLS	POPTWRLLAM	LAGIAYFFLE	NFLNL-----LP 400	
QY	459	ENDDDV-----	IKKOLSKYES	QSLSTNEEK	VTDDORTGYLRADSOEP 501	
DB	401	RDPEDLBDG	PCGSHSHSGH	SGVSLQAP	S-----LR--QPKP 439	
QY	502	SHFDSQOPAV	LEEEVMTAHAP	QEVNEYVY	PGCKNKHSHFDTL	QGSDLLIHHDY 561
DB	440	PHESGRAD	IVAE-----	-----	-----SPELL-----457	
QY	562	HHILHHHHH	QHHPHSHQRY	SREELKD	GVATLAWNVIMG	DGLHNPSDGLATCAFTG 621
DB	458	-----NPEP	-----	-----	-----	-----
QY	622	LSSGLSTSV	AVFCHELPH	ELGDFAVLL	KAGMTVKQAV	LYNALSMALAYLGM 672
DB	498	WKTGLAP	SLAVFCHELPH	ELGDFAAAL	HAGLSVRO	ALLINLASALTAFAGL 548

RESULT 9

```

US-10-017-161-2394
; Sequence 2394, Application US/10017161
; Publication No. US2003014368A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKITAMA, YUTAKA
; APPLICANT: ASURAMANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-C
; FILE REFERENCE: 08435/0152
; CURRENT APPLICATION NUMBER: US/10/017
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001-2461

```

```

; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2394
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-2394

Query Match          10.0%; Score 401.5; DB 12; Length 737;
Best Local Similarity 23.3%; Pred. No. 3.2e-26;
Matches 165; Conservative 95; Mismatches 205; Indels 243; Gaps 25

QY 130 SHHNRASGNKKKALCPD-----HDSUSSGKDPNS-----QKGARPEHAS 173
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 121 SHADHLLALLESKALTGSLWLLQRMARAAGOTPRMACVDIPOLLSEAVGAGAGSAG 180
QY 174 GRRNVKDSVSASVTSTVNTVNTSEGTHTLETITPRPGKLFP-KDVSSTPPSVTGSRV 232
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 181 G-----VLAALLDHVRSQCF-HALPSQVYFDFVFQHSSEVMTLAASTSL 227
QY 233 SRLAGRTNE-----SVSPRKGFMYSNNW 258
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 228 GGRAGATDSRCQLFSLPELSALMQLRGREAHSDHSHRGASRRDPVPLISSNSS 287
QY 259 ENPQECFNASKLTSHGMGIOVPLNATEFNVLCPALINGOIDARSCLHTSEKKAIEPKT 318
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 268 VMDIVCLSDRYAAALGSEQGYTPPAWQLGFPALLQQLSGAC--TSQSR--PPVQ 341
QY 319 YSLQIA--WVGGFIAIISIFLLGVILVPLMN--RVYFKFLLSFLVALAYTGLSGDAFL 375
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 342 DQLSQSERLYGLSLATLLICLCAVFGILLTCTGCRGWTHYTIQTLFLSLAVGAVTGDAVL 401
QY 376 HLLPHSHASHHSHSHEEPAMEMKGPLFS-----HLSSQNTSESAYFDSTWKG--L 425
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 402 HLTPT-----KVCPHKPATALSHSPWTPRPCWGCIIHAKRAHSP-----PGALL 447
QY 426 TALGGLYFMFLVSHVTLTIQKFKDKKKKNOKKPENDDVEIKKOLSKYESOLSTNEEKVD 485
    || | | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 448 AMLAGLYAFFLENF-----NLKQPRDPDLF----- 475
QY 486 TDDRTEGYLRADSQSFHSFDSQPAVLEEEVMTAHAFQEVYNYVPGCKNKCCHSEPH 545
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 476 --DQPCGSHSHSGHSGHVSQLA-----PSELQPKPP-----H 509
QY 546 DTLGQSDDLTHHEHDYHLLHHHHQNHHP-HHSQR-----YSREELK 588
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 510 E-GSRADIV-----SGRQMPHPTRSPSHRPLPTLPAPQACGSASRE--- 552
QY 589 DAGVATYLAWVINGDLGNFSDGLAIGAFTGSLSTSVAYTCHPEL----- 641
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 553 ---LRLPYMTLGDVAVHNFADGLAVGAFASSWKTLGTLATSLAVTCHPELGERRGL 609
QY 642 -----GQFVALLKAGMTVRQAV 658
    || : | : | | | : | : | : | : | : | : | : | : | : | : | : |
Db 610 EGDGRTGLGGETWNRWAGPCRWAGPHLAGWDRKQVGRGLEGDFALLHAGLSVRQAL 669
QY 659 LYNLSAMLAYLGMATGIIGHYAENVSMWIFALTAGLFMYVALVDMV 706
    || : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 670 LLNLASALTAFAGLYVALAVG-VSESEAWILAVATGLFLYVALCDMV 716

RESULT 10
US-10-102-806-743
; Sequence 743, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103PCL
; CURRENT APPLICATION NUMBER: US/10/102.806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298

```



```

; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/069,957
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/074,121
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/081,563
; PRIOR FILING DATE: 1998-04-13
; PRIOR APPLICATION NUMBER: 60/096,116
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/099,273
; PRIOR FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: Patent.pm
; SEQ ID NO 196
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-319-763-196

Query Match
Best Local Similarity 28.7%; Pred. No. 1.4e-06;
Matches 43; Conservative 27; Mismatches 68; Indels 12; Gaps 4;

QY 1 MARKSVILITFALSVNPLHKLAAFPQTTEK-----ISPNWESGINVDLAISTR 53
      | | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 MHTKFCILCLTTFPHCHNCHE-EHDHGPEALHRQHRGWTLEP---SKFSQAANEK 60

QY 54 QYHLOQLFRYGENNLSVEGFRKLLQNTGIDKIKYRIHHDDHSDHSDHSDHSDHSD 113
      | : : : | | | | | | | | | | | | | | | | | | | | | | |
Db 61 KYIEKLFYRGNGRLSFFGLKLLTNLGLGRKYVEINHEDLGH-DHVSHLGILAVGE 119

QY 114 HEHSDHEHSDHSDHSHNHAASGNKRK 143
      : | | | | | | | | | | | | | | | | | | | | | |
Db 120 GKHFHSHNQSHHNLSENQTVISVSTKK 149

RESULT 15
US-09-946-374-95
; Sequence 95, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, F. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830PIC1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099815
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100388
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100584
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100661
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100662
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100664
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100683
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100684
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100710
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100711
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100849
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/100930
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101014
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101068
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101071
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/101279
```

; PRIOR FILING DATE: 1998-09-22
; PRIOR APPLICATION NUMBER: 60/101471
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101472
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101474
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101475
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101476
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101479
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101743
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101915
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/102207
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102240
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102307
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102330
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102331
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: 60/102484
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102487
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102570
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102571
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: 60/102684
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102687
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/102965
; PRIOR FILING DATE: 1998-10-02
; PRIOR APPLICATION NUMBER: 60/103258
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103314
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103315
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103328
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103395
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103396
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103401
; PRIOR FILING DATE: 1998-10-07
; PRIOR APPLICATION NUMBER: 60/103449
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: 60/103633
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103678
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103679
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/103711
; PRIOR FILING DATE: 1998-10-08

; PRIOR APPLICATION NUMBER: 60/104257
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 60/104987
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105000
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105002
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: 60/105104
; PRIOR FILING DATE: 1998-10-21
; PRIOR APPLICATION NUMBER: 60/105169
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105266
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: 60/105693
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105694
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: 60/105807

Query Match 4.1%; Score 166; DB 11; Length 307;
Best Local Similarity 24.8%; Pred. No. 4.1e-06;
Matches 68; Conservative 40; Mismatches 80; Indels 86; Gaps 12;

Qy 532 VPRGCKNCHSHFHDHIGOSDOLHHHHYH-----ILHHHHHQNHPHSH- 578
Db 55 VPEG-----VHALYEDILEGK---HQASETHNVIASDKAEKSVVHEHEHSHDTQLHA 106
Qy 579 -----SQRYSRBELKDAGVATLAWMTMGDLHNFTSDGLAIG-A 616
Db 107 YIGVSLVLGFVFMILLVDQIGNSHVHSTDDPEARSNSKITTTGLVWVHAAADGVALGAA 166
Qy 617 AFTEGLSSGLSTSVAVFCHPELPHGLDFAVLKAGM---TVKQAVLYNALS-----MLAY 669
Db 167 ASTSTSVOLIVFVATMLKAPAAAGLVSLFMHAGLERNIRKHLVLFALAAPVMSWTY 226
Qy 670 LGM-----ATGIFIGHYAENYSMMIFALTAGLFMYVALVDVPE---MLHND 714
Db 227 LGLSKSKSEALSEVNATGV-----AMLSAGTFLYVATVHVLPVEVGIGHSHK 274
Qy 715 SD-----HCCSRWGYFFLQAGMLGLFGIMLLISI 744
Db 275 PDATGGRGLSR-----LEVALVLGCLPLISV 303

Search completed: September 12, 2003, 22:27:24
Job time : 65 secs